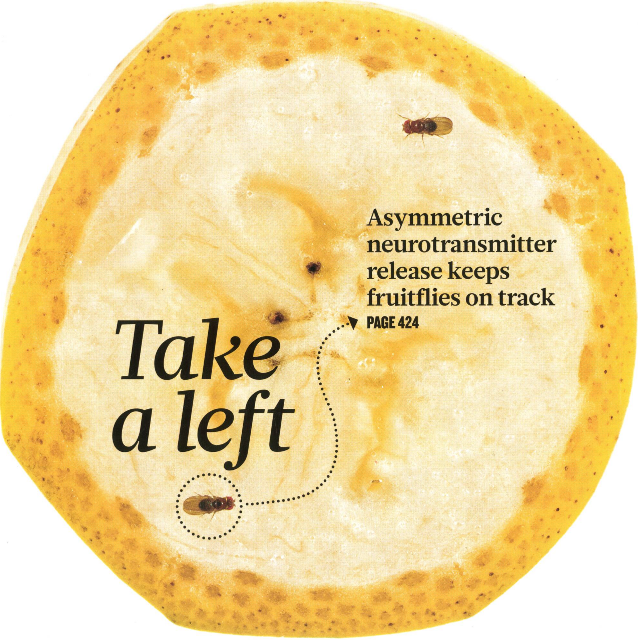
INSIGHT
Frontiers in biology

nature

THE INTERNATIONAL WEEKLY JOURNAL OF SCIENCE



Take a left

Asymmetric neurotransmitter release keeps fruitflies on track

PAGE 424

MENTAL HEALTH

FUKUSHIMA LEGACY

Nuclear meltdown left psychological scars

PAGE 290

CULTURE

SCREEN TEST

Science at the Sundance Film Festival

PAGE 302

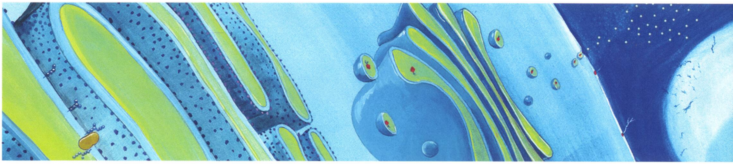
NANOSCALE MATERIALS

DIAMOND LEAGUE

Ultrahard cubic boron nitride

PAGE 385

➲ **NATURE.COM/NATURE**
17 January 2013 £10
Vol. 493, No. 7432

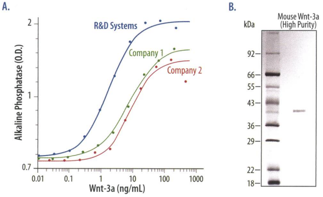


R&D Systems Cytokines & Growth Factors

Setting the Standard for the Life Science Industry for over 25 Years

- ✓ R&D Systems currently offers more than 3,000 bioactive recombinant proteins
- ✓ We are the most referenced manufacturer of recombinant & natural proteins
- ✓ Our scientists are dedicated to developing the highest quality products & new processes that meet the needs of researchers

High purity proteins, traditional & cell culture grade • Animal-free proteins • GMP proteins



A. Activity Comparison Data for Recombinant Mouse Wnt-3a. The MC3T3-E1 mouse preosteoblast cell line was treated with increasing concentrations of R&D Systems Recombinant Mouse Wnt-3a (Catalog # 1324-WN) or with mouse Wnt-3a from two other companies. The R&D Systems protein shows greater than a four-fold increase in activity compared to the other commercially available proteins.

B. SDS-PAGE Analysis of High Purity Recombinant Mouse Wnt-3a. High Purity Recombinant Mouse Wnt-3a (Catalog # 1324-WNP; 1 µg/lane) was loaded on a 12% SDS-PAGE gel under reducing conditions and visualized by silver staining.

www.RnDSystems.com/CytokinesGrowthFactors

R&D Systems Tools for Cell Biology Research™

For research use only. Not for use in diagnostic procedures.

R&D SYSTEMS®

I seek the future.
Welcome to MiSeq Control Software
MiSeq
illumina

MiSeq® Next-generation sequencing for all you seek.

You want amazing accuracy and performance on your benchtop. Illumina delivers—yet again. The MiSeq personal sequencer is the only fully integrated, truly end-to-end benchtop solution around. It's just one more example of why Illumina solutions generate a remarkable 90% of all the world's sequencing data. Discover what's possible.

www.illumina.com/miseq

illumina®

MACMILLAN
SCIENCE COMMUNICATION

Scientific Writing and Publishing Workshops

Nanjing Agricultural University

and

Macmillan Science Communication (MSC),
an exclusive partner of Nature Publishing Group,

are holding an interactive Scientific Writing and Publishing Workshop in Nanjing
19th and 20th February 2013

- The MSC intensive two-day workshop will train 30 researchers with previous experience of publishing research papers, how to prepare and publish high quality research papers in international, peer-reviewed journals.
- Training will be led by *Nature Genetics* Chief Editor, Dr. Myles Axton, and former *Nature* and *Nature Genetics* Editor, Dr. Wayne Peng, who bring their expertise and editorial insight from working at a prestigious scientific publishing house.
- The workshop will cover the writing and publishing process from start to finish, using short lectures punctuated with practical exercises and group discussions that focus on the publishing issues that researchers face today.
- On the 21st February, MSC will provide independent feedback on pre-selected, unpublished papers written by workshop attendees in face to face consultations.

This workshop is brought to you by Macmillan Science Communication, exclusive partner of Nature Publishing Group, publisher of *Nature*.

Nanjing Agricultural University | **www.njau.edu.cn**
MSC workshops | **www.macmillanscience.com/training**

nature publishing group npg

CONTENTS

17 January 2013 / Vol 493 / Issue No 7432

THIS WEEK

EDITORIALS

271 **MENTAL HEALTH**
Troubling thoughts
Fukushima Health Survey could benefit future disaster survivors worldwide

271 **INSIDER TRADING**
Knowledge trades
Researchers must be careful not to divulge confidential information

272 **GENETICS**
Natural history
Genetic studies of wild behaviour should initially focus on select species

WORLD VIEW

273 **China has the capacity to lead in carbon trading**
Qiang Wang
If it can guarantee transparent data collection, China would make a fertile ground for a carbon market

RESEARCH HIGHLIGHTS

274 **SELECTIONS FROM THE SCIENTIFIC LITERATURE**
Robot ribosome / Out of Australia / Hair-cell replacement / Predator avoidance / Global-warming estimates accurate / Wasps aren't useless

SEVEN DAYS

276 **THE NEWS IN BRIEF**
Europe to launch food-safety initiative / Antarctic team prepares to drill into Lake Whillans / Global investment in clean energy drops off

CAREERS

441 **EDUCATION**
Free-range learning
Informal science education is gaining ground despite a struggle for funding

443 **Q&A**
Changes at the NIH
Francis Collins explains changes in the NIH's approach to career funding

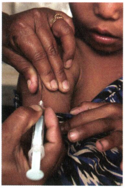
NEWS IN FOCUS

279 **AFRICA**
Plans for clinical trial of antibody against HIV draw fire

280 **FINANCE**
Universities wrestle with the ethics of expert networks

282 **STEM CELLS**
US court rules in favour of research using human embryonic stem cells

283 **TECHNOLOGY**
Quantum dots prepare for their small-screen debut

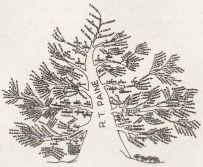


284 **GENETICS**
DNA analysis provides clues to complexities of burrowing

285 **HEALTH**
WHO backs use of inactivated vaccine in final push to eradicate polio

FEATURES

290 **NUCLEAR SAFETY**
Fallout of fear
Nearly two years on from the Fukushima disaster, the psychological scars remain



SCIENTIFIC FAMILIES
Dynasty
How Bob Paine influenced an academic family that changed the face of ecology **PAGE 286**

COMMENT

ENVIRONMENT
Model all life on Earth
Ecologists should follow climatology's lead and simulate the globe **PAGE 295**

298 **POLICY**
Two years of blogging the NIH
Sally Rockey
An NIH insider's popular blog sets standards in transparency

BOOKS & ARTS

300 **CONSERVATION**
The commonness of rarity
Stuart Pimm

301 **STATISTICS**
Number-crunching in the raw
David Vaux

302 **FILM**
Walking the edge at Sundance
Jascha Hoffman

CORRESPONDENCE

304 Crowding-out true science / Cancer researchers stick to 2020 target / Whose carbon? / Journal objectivity / Sleeping-pill side effects / Lost phylogenetic trees / Marked amphibians

OBITUARY

306 **Rita Levi-Montalcini (1909–2012)**
Ralph A Bradshaw

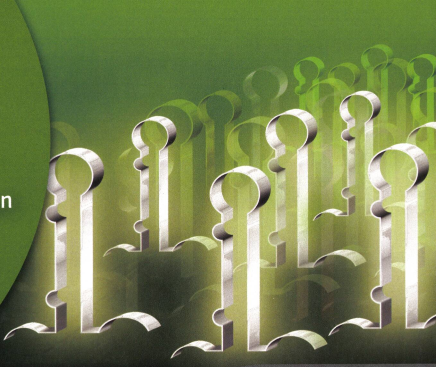
FUTURES

446 **Eros for Annabelle**
Ian Whates

NEW

PLATINUM *Select* shRNA-mir

Groundbreaking knockdown performance with shERWOOD algorithm designed hairpins



Find shRNA-mir products targeting your gene of interest using

FETCH my *gene*™

www.transomic.com

Developed by Dr. Greg Hannon and colleagues at Cold Spring Harbor Laboratory (CSHL), PLATINUM *Select* shRNA-mir are designed using the new shRNA-specific shERWOOD algorithm developed using data generated by functionally testing over 250,000 shRNA sequences in a high-throughput sensor assay. Targeting human and mouse genomes, these new shRNA-mir libraries offer dramatically improved knockdown performance even when represented at single copy in the genome.

- ***Powerful knockdown performance*** - *advances RNAi tools*
- ***Enhanced specificity and reduced toxicity*** - *with microRNA scaffold*
- ***100% Guaranteed knockdown*** - *All shRNA-mir targeting your gene will produce ≥ 70% knockdown at the mRNA level*

NEW

Epigenetics focused shRNA-mir libraries -
Available in arrayed or pooled screening formats.

transOMIC technologies™

www.transomic.com

Enabling Discovery Across the Genome

CONTENTS

17 January 2013 / Vol 493 / Issue No 7432

RESEARCH

NEW ONLINE

307 Papers published this week at nature.com

NEWS & VIEWS

308 EVOLUTION
A history of give and take
Foraminifera fossils link extinctions with environmental change
Steven M Holland **SEE LETTER P.398**

309 EARTH SCIENCE
Plumbing nickel from the core
A potential new mantle source
Michael J Walter **SEE LETTER P.393**

310 STEM CELLS
Surf the waves of reprogramming
A molecular roadmap of cellular reprogramming to iPS cells
Ignacio Sancho-Martinez & Juan Carlos Izpisua Belmonte

312 EVOLUTIONARY GENETICS
Genes for home-building
Genetic modules responsible for burrowing style in oldfield mice
Patrick Goymer **SEE LETTER P.402**

312 NEUROSCIENCE
Memory and the single molecule
The protein PKM-ζ is less important in maintaining memory than thought
Paul W Frankland & Sheena A Josselyn
SEE LETTERS P.416 & P.420

313 ANALYTICAL CHEMISTRY
Defects visualized in porous solids
Confocal fluorescence microscopy of metal–organic frameworks
Christian Serre

314 COMPUTATIONAL MATERIALS SCIENCE
Trustworthy predictions
A method to precisely compute the properties of certain solids
Paul R C Kent **SEE ARTICLE P.365**

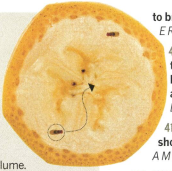
ARTICLES

365 MATERIALS SCIENCE **Towards an exact description of electronic wavefunctions in real solids**
G H Booth, A Grüneis, G Kresse & A Alavi **SEE N&V P.314**

371 NEUROSCIENCE **Autism-related deficits via dysregulated eIF4E-dependent translational control**
C G Gkogkas et al.

ON THE COVER

Taking turns



A fruitfly knows right from left when confronted with an asymmetric odour plume. A new study demonstrates a neuronal basis for that ability which makes use of the fact that central neurons on the fly's right spike a few milliseconds before those on the left.
PAGE 424

LETTERS

378 ASTRONOMY **Pulsed accretion in a variable protostar**
J Muzerolle, E Furlan, K Flaherty, Z Balog & R Gutermuth

381 ASTRONOMY **Planetary system disruption by Galactic perturbations to wide binary stars**
N A Kaib, S N Raymond & M Duncan

385 MATERIALS SCIENCE **Ultrahard nanotwinned cubic boron nitride**
Y Tian et al.

389 CLIMATE SCIENCE **Multidecadal variability in East African hydroclimate controlled by the Indian Ocean**
J E Tierney, J E Smerdon, K J Anchukaitis & R Seager

393 EARTH SCIENCE **Nickel and helium evidence for melt above the core–mantle boundary**
C Herzberg et al. **SEE N&V P.309**

398 PALAEONTOLOGY **Oceanographic controls on the diversity and extinction of planktonic foraminifera**
S E Peters, D C Kelly & A J Fraass
SEE N&V P.308

402 GENETICS **Discrete genetic modules are responsible for complex burrow evolution in *Peromyscus* mice**
J N Weber, B K Peterson & H E Hoekstra **SEE N&V P.312**

406 CANCER **Mosaic *PPM1D* mutations are associated with predisposition to breast and ovarian cancer**
E Ruark et al.

411 NEUROSCIENCE **Exaggerated translation causes synaptic and behavioural aberrations associated with autism**
E Santini et al.

416 NEUROSCIENCE ***Prkcz* null mice show normal learning and memory**
A M Lee et al. **SEE N&V P.312**

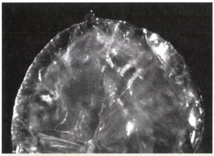
420 NEUROSCIENCE **PKM-ζ is not required for hippocampal synaptic plasticity, learning and memory**
L J Volk, J L Bachman, R Johnson, Y Yu & R L Huganir **SEE N&V P.312**

424 NEUROSCIENCE **Asymmetric neurotransmitter release enables rapid odour lateralization in *Drosophila***
Q Gaudry, E J Hong, J Kain, B L de Bivort & R I Wilson

429 MICROBIOLOGY **Bacteriophage genes that inactivate the CRISPR/Cas bacterial immune system**
J Bondy-Denomy, A Pawluk, K L Maxwell & A R Davidson

433 STEM CELLS **Mammalian heart renewal by pre-existing cardiomyocytes**
S E Senyo et al.

437 STRUCTURAL BIOLOGY **Structure and function of the initially transcribing RNA polymerase II–TFIIB complex**
S Sainsbury, J Niesser & P Cramer



MATERIALS SCIENCE

Hard times

Cubic boron nitride now rivals diamond in its ultrahard state. **PAGE 385**

KELLY KRAUSE/NATURE — IMAGES BY SYLVIE BOUCHARD, DARIUSZ MAJGIER/SHUTTERSTOCK

nature.com marketplace

Do you want to save money on your lab costs?

Receive exclusive discounts from top suppliers of lab products and keep up-to-date with the latest product information as published in *Nature* and *Nature Methods*.

Nature.com marketplace is the new product focused website brought to you by Nature Publishing Group.

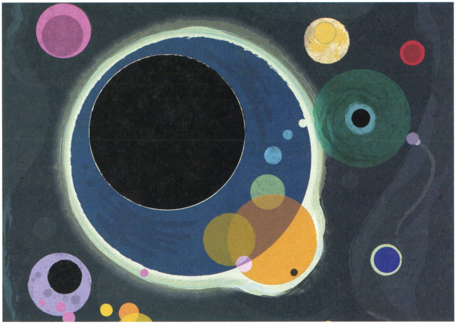
Visit nature.com/marketplace

nature publishing group npg

INSIGHT

17 January 2013 / Vol 493 / Issue No 7432

FRONTIERS IN BIOLOGY



NIK SPENCER/NATURE

REVIEWS

318 Cytosystems dynamics in self-organization of tissue architecture
Our knowledge of the principles that determine the shape and size of organs in the body is still limited. Organogenesis is driven by various cellular communications, and understanding the role of these in a dynamic system requires an understanding of their context in time and space. Over the past few years, new approaches, such as three-dimensional stem-cell cultures to mimic the self-organization of cells to form tissues, have made analysing these complex dynamics easier. Emerging technologies and technical improvements, including four-dimensional measurement, theoretical modelling and experimental reconstitution, may help our understanding of collective cell behaviours in organ formation.
Yoshiki Sasai

327 Activity-dependent neuronal signalling and autism spectrum disorder
Recent evidence suggests that many of the genes implicated in autism spectrum disorder encode proteins that are central to the activity-dependent pathways that regulate synaptic development and plasticity. Sensory experience promotes the development of inhibitory synapses during the first few years of life — a time when most of the features of the disorder begin to manifest. By investigating the interaction between autism-associated genes and activity-dependent signalling networks, we can gain insight into the pathogenesis of autism spectrum disorder.
Daniel H. Ebert & Michael E. Greenberg

338 mTOR is a key modulator of ageing and age-related disease
Having the tools to slow ageing in humans could be just a matter of time. Targeting mTOR as part of the nutrient response pathway can extend lifespan in several model organisms, and can protect against a variety of age-related diseases in mammals. Inhibitors of this pathway are already clinically approved for some diseases, with others in development. Many questions still need to be answered about the risks and side effects that are associated with mTOR inhibitors, but they may one day be used to slow ageing and protect against age-related diseases in humans.
Simon C. Johnson, Peter S. Rabinovitch & Matt Kaeberlein

346 Metabolism of inflammation limited by AMPK and pseudo-starvation
Altered metabolism could be a key signal for both inflammation and the development of tumours. Although our understanding of the changes that occur in inflammatory cells has lagged behind that of tumour cells, it is thought that they may share many of the same features. Cells that participate in inflammation can have higher glucose uptake and glycolysis, and an increase in pentose phosphate pathway activity. The transcription factor HIF-1α and AMP-activated protein kinase (AMPK) are proposed to have opposing roles in these metabolic changes. By contrast, anti-inflammatory cells have lower glycolytic rates and an increased rate of oxidative metabolism. AMPK is activated by some anti-inflammatory agents, which may induce a state similar to pseudo-starvation.
Luke A. J. O'Neill & D. Grahame Hardie

356 Fanconi anaemia and the repair of Watson and Crick DNA crosslinks
Fanconi anaemia is a rare disorder characterized by profound genomic instability and sensitivity to DNA damage. Fifteen genes have been identified as mutated in patients with the disorder, and all 15 gene products act together in a pathway dedicated to the repair of DNA interstrand crosslinks. Investigating how crosslink-repair proteins interact with the DNA repair pathways that act on other types of lesions is valuable for creating a picture of how human cells cope with genotoxic insults. By using mice deficient in Fanconi anaemia genes, researchers have examined how the proteins that the genes encode affect development and protect against DNA damage. Understanding the consequences of this disorder at the molecular and organismal level will provide insight into the specific mechanics of DNA repair in the cell and into the pathogenesis and potential treatment of cancer in general.
Molly C. Kottemann & Agata Smogorzewska

Manpei Suzuki Diabetes Foundation

Calling for 2013 nominations for Manpei Suzuki International Prize for Diabetes Research

Go to www.manpeisuzukiprize.org

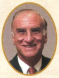
Deadline: July 1, 2013

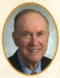
- The Manpei Suzuki Diabetes Foundation is pleased to call for a nomination of a researcher who has accomplished creative, innovative researches of highest quality in Diabetes Research.
- Award comprises of a Certificate of Honor, US$150,000 and a Japanese objet d'art.
- A member of ADA, EASD, JDS or other scientific societies can submit one nomination yearly.

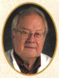
2012 Winner

Graeme I. Bell, PhD

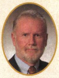
Louis Block Distinguished Service Professor in
Medicine and Human Genetics
The University of Chicago

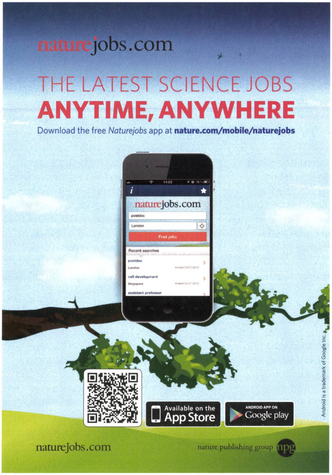

2008
C. R. Kahn, MD


2009
D. F. Steiner, MD


2010
C. E. Mogensen, MD

2011
M. Kasuga, MD, PhD


2012
G. I. Bell, PhD


naturejobs.com
THE LATEST SCIENCE JOBS
ANYTIME, ANYWHERE
Download the free Naturejobs app at nature.com/mobile/naturejobs
naturejobs.com
postdoc
London
Find jobs
Recent searches
postdoc
cell development
assistant professor
Available on the App Store
ANDROID APP ON Google play
Android is a trademark of Google Inc.
naturejobs.com
nature publishing group npg

nature

17 January 2013 / Vol 493 / Issue No. 7432

EDITORIAL

LONDON, MUNICH, WASHINGTON DC, NEW YORK, BOSTON, PARIS, SAN FRANCISCO, TOKYO **nature@nature.com**

EDITOR-IN-CHIEF: Philip Campbell
EXECUTIVE EDITOR: Véronique Kiermer
CHIEF MAGAZINE EDITOR: Tim Appenzeller
THIS WEEK/NEWS/FEATURES: Alison Abbott, David Adam, Ananyo Bhattacharya, Geoff Brumfiel, Declan Butler, Ewen Callaway, Davide Castelvecchi, Erika Check Hayden, Daniel Cressey, David Cyranoski, Natasha Gilbert, Eric Hand, Heidi Ledford, Corie Lok, Brendan Maher, Emma Marris, Richard Monastersky, Helen Pearson, Eugenie Samuel Reich, Quirin Schiermeier, Jeff Tollefson, Richard Van Noorden, Meredith Wadman, M Mitchell Waldrop **COMMENT:** Sara Abdulla, Joanne Baker, Rosalind Cotter, Barbara Kiser, Alison McCook, Lucy Odling-Smee **NATURE PODCAST/VIDEO:** Adam Rutherford, Noah Baker, Thea Cunningham, Kerri Smith, Charlotte Stoddart **NEWS AND VIEWS:** Sadaf Shadan, Ana Lopes, Andrew Mitchinson, Marian Turner **RESEARCH:** Ritu Dhand, Karl Ziemelis, Francesca Cesari, I-han Chou, Tanguy Chouard, Rosamund Daw, Vera Domingues, Alex Eccleston, Angela Eggleston, Joshua Finkelstein, Henry Gee, Patrick Goymer, Noah Gray, Marie-Thérèse Heemels, Magdalena Helmer, Karen Howell, Andrew Jermy, Barbara Marte, Andrew Mitchinson, Juliane Mössinger, Deepa Nath, Leslie Sage, Magdalena Skipper, Clare Thomas, Maria Trajkovska, John VanDecar, Liesbeth Venema, Ursula Weiss, Michael White **INSIGHTS/REVIEWS/PERSPECTIVES:** Ursula Weiss **CAREERS:** Gene Russo, Karen Kaplan **TECHNOLOGY FEATURES:** Vivien Marx **FUTURES:** Colin Sullivan **SUPPLEMENTS:** Herb Brody, Michelle Grayson, Tony Scully **SUBEDITORS:** Colin Sullivan, Mary Abraham, Sarah Archibald, Nicola Bailey, Nicolas Fanget, Isobel Flanagan, Paul Fletcher, Anne Haggart, Sasha Lewis, Dinah Loon, Anna Novitzky, Yukie Ozawa, Katherine Perry, David Price, Jenny Rooke, David Winters, Anna York **PRODUCTION:** Jenny Henderson, Leonora Dawson-Bowling, Dan Foley, Stewart Fraser, Susan Gray, Simon Gribbin, Hazel Mayhew, James McQuat, Emilia Orviss, Yvonne Strong, Robert Sullivan, Roberta Tarran, Charles Wenz **ART AND DESIGN:** Kelly Buckheit Krause, Wesley Fernandes, Madeline Hutchinson, Barbara Izdebska, Paul Jackman, Jasiek Krzysztofiak, Annthea Lewis, Chris Ryan, Nik Spencer, Claire Welsh, Mark Wood **JAPAN EDITION:** Eiji Matsuda, Mika Ishida, Mitsuyo Utsugi, Sayaka Ayame **PRESS OFFICE:** Alice Henchley, Neda Afsarmanesh, Lisa Boucher, Jenny Chapman, Rachel Twinn, Rebecca Walton **ADMINISTRATION:** Paul Sng, Roseann Campbell, Peter Cary, Ffion Cleverley, Dana Coles, Ashlee Gagui, Peter Hewitt, Caroline McLean, Anastasia Panoutsou, Kenneth Simpson, Kimberly Tolleson, Michelle Viscuso, Diane Yorke

PUBLISHING

LONDON **feedback@nature.com**
MANAGING DIRECTOR & PUBLISHER: Steven Inchcoombe
ASSOCIATE PUBLISHER: Arash Hejazi
PUBLISHING PROJECT MANAGERS: Claudia Deasy, Christian Manco
TOKYO **feedback@natureasia.com**
REGIONAL MANAGING DIRECTOR — ASIA-PACIFIC: David Swinbanks
DIRECTOR — ASIA-PACIFIC: Antoine E Bocquet
OPERATIONS DIRECTOR — ASIA-PACIFIC: Hiroshi Minemura

DISPLAY ADVERTISING, SPONSORSHIP & NATUREJOBS

GLOBAL HEAD: Andrew Douglas a.douglas@nature.com

DISPLAY ADVERTISING AND SPONSORSHIP

GLOBAL HEAD: Gerard Preston g.preston@nature.com
HEAD OF SPONSORSHIP: Reya Silao Tel: +44 (0)20 7843 4977 (r.silao@nature.com)
EUROPE/REST OF WORLD display@nature.com Tel: +44 (0)20 7843 4960
GERMANY/SWITZERLAND/AUSTRIA: Sabine Hugi-Fürst Tel: +41 52761 3386
UK/IRELAND/SCANDINAVIA/SPAIN/PORTUGAL: Evelina Rubio Hakansson Tel: +44 (0)20 7014 4079
FRANCE/BELGIUM/THE NETHERLANDS/LUXEMBOURG/ITALY/OTHER EUROPE: Nicola Wright Tel: +44 (0)20 7843 4959
NORTH AMERICA **display@natureny.com**
NEW ENGLAND/EASTERN CANADA: Sheila Reardon Tel: +1 617 475 9202
NEW YORK/MID-ATLANTIC/SOUTHEAST: Jim Breault Tel: +1 212 726 9334
MIDWEST: Mike Rossi Tel: +1 212 726 9255
WEST COAST/WESTERN CANADA: George Lui Tel: +1 415 781 3804
NORTHWEST: Kayla McCutchan Tel: +1 415 403 9038; Fax: +1 415 781 3805
ASIA-PACIFIC **display@natureasia.com**
ASIA-PACIFIC HEAD: Kate Yoneyama
JAPAN: Ken Mikami Tel: +81 3 3267 8765
GREATER CHINA/SINGAPORE: Gloria To Tel: +852 2811 7191
INDIA: Vikas Chawla Tel: +91 124 288 1057

SPONSORSHIP

sponsorship@nature.com
SENIOR BUSINESS DEVELOPMENT MANAGER: David Bagshaw Tel: +1 212 726 9215
BUSINESS DEVELOPMENT MANAGER: Greg Valero Tel: +1 212 726 9360
BUSINESS DEVELOPMENT MANAGER: Will Piper Tel: +44 (0)20 7014 4181

NATUREJOBS

naturejobs@nature.com
EUROPEAN HEAD OFFICE, LONDON: Tel: +44 (0)20 7843 4961
GLOBAL HEAD OF CLASSIFIED ADVERTISING: Nils Moeller (n.moeller@nature.com)
BUSINESS DEVELOPMENT MANAGER, MIDDLE EAST: Jon Giuliani (j.giuliani@nature.com)
ADVERTISING PRODUCTION MANAGER: Stephen Russell **NATUREJOBS ONLINE PRODUCTION:** Dennis Chu
US HEAD OFFICE, NEW YORK: naturejobs@us.nature.com **US SALES MANAGER:** Ken Finnegan
JAPAN HEAD OFFICE, TOKYO: Tel: +81 3 3267 8765 **ASIA-PACIFIC HEAD:** Kate Yoneyama
ASIA-PACIFIC SALES & BUSINESS DEVELOPMENT MANAGER: Yuki Fujiwara
BUSINESS DEVELOPMENT MANAGER, GREATER CHINA, SINGAPORE: Gloria To Tel: +852 2811 7191
INDIA: Vikas Chawla Tel: +91 1242881057 (v.chawla@nature.com)

MARKETING & SUBSCRIPTIONS

USA/CANADA/LATIN AMERICA **subscriptions@us.nature.com**
Tel: (USA/Canada)+1 866 363 7860; (outside USA/Canada) +1 212 726 9223
MARKETING: Sara Girard **FULFILMENT:** Karen Dawson

JAPAN/CHINA/KOREA **subscriptions@natureasia.com**
Tel: +81 3 3267 8751 **MARKETING:** Sachiyo Ikeda **FULFILMENT:** Masayo Tanaka

EUROPE/REST OF WORLD **subscriptions@nature.com**
Tel: +44 (0)1256 329242
MARKETING: Elena Woodstock, Marcus Parker, Hannah Phipps **FULFILMENT:** Jean Burrows

INDIA **npgindia@nature.com**
Tel: +91 124 2881053/54
HEAD OF BUSINESS DEVELOPMENT, INDIA: Debashish Brahmachari **MARKETING:** Harpal Singh Gill

ANNUAL SUBSCRIPTIONS (UK/Europe, including post and packing)
FULL RATE: UK £1,947; Europe €3,014 **PERSONAL RATES:** UK £135; Europe €209; Rest of World £210
STUDENT RATES: UK £79; Europe €122; Rest of World £125
Individual rates apply to payments by personal cheque or credit card. Back issues: £20.00.

ANNUAL SUBSCRIPTIONS (USA, including post and packing)
INSTITUTIONAL/CORPORATE RATE: $3,790; **PERSONAL RATE:** $199; **STUDENT RATE:** $99; **POSTDOC RATE:** $119
Printed in USA. Individual rates available only to subscribers paying by personal check or credit card. Orders for student/postdoc subscriptions must be accompanied by a copy of student ID. Rates apply to USA, Canada, Mexico/Central & South America. Add 7% GST tax in Canada (Canadian GST number 140911595). **BACK ISSUES:** US$20.00.

ANNUAL SUBSCRIPTIONS (Asia, including post and packing)
JAPAN: **年間購読料金：** *Nature* のみ 536,300 円 +26,815(消費税)=563,115円(合計)
Nature + *Nature Digest* 574,500 円 +28,725(消費税)=603,225円(合計)
お問い合わせ：日本出版貿易株式会社 商品統括部 雑誌営業課
〒101-0064 東京都千代田区猿楽町1-2-1 Tel: 03-3292-3767 Fax: 03-3292-8766
個人/学生購読について： ご本人名義で個人的に支払い、私物として購読する場合に適用されます。
NPG ネイチャー アジア・パシフィック フルフィルメント
〒162-0843 東京都新宿区市谷田町2-37 千代田ビル Tel: 03-3267-8751Fax: 03-3267-8752
CHINA: www.natureasia.com/ch/subscribe; **KOREA:** www.natureasia.com/korea/subscribe

NATURE ONLINE

WWW.NATURE.COM/NATURE

CHIEF TECHNOLOGY OFFICER: Howard Ratner **ASSOCIATE DIRECTOR, NATURE.COM:** Daniel Pollock **PLATFORM SERVICES:** Jeremy Macdonald **WEB PRODUCTION:** Alexander Thurrell, Dipti Shah **WEB DESIGN:** Tom Scott, Juan-Carlos Sobrino, Luke Stavenhagen, Cianna Timbers **PLATFORM INTERFACES:** Amanda Ward

SITE LICENCES, FULFILMENT & CUSTOMER SERVICES

SITE LICENCES: npg.nature.com/libraries feedback@nature.com
CHIEF OPERATIONS OFFICER: John Carroll **ASSOCIATE DIRECTOR OF OPERATIONS:** Dominic Pettit
CUSTOMER SERVICE: Liz Davidson **ORDER PROCESSING OPERATIONS:** Steve Poulter

NATURE PUBLISHING GROUP OFFICES

LONDON: The Macmillan Building, 4 Crinan Street, London N1 9XW Tel: +44 (0)20 7833 4000 Fax: +44 (0)20 7843 4596/7
BASINGSTOKE: Nature Publishing Group, Subscriptions, Brunel Road, Basingstoke, Hants RG21 6XS, UK Tel: +44 (0)1256 329242
WASHINGTON DC: 968 National Press Building, 529 14th St NW, Washington DC 20045-1938 Tel: +1 202 737 2355
NEW YORK: 75 Varick St, 9th Floor, New York, NY 10013-1917 Tel: +1 212 726 9200
BOSTON: 25 First Street, Suite 104, Cambridge, MA 02141 Tel: +1 617 475 9275
SAN FRANCISCO: 225 Bush Street, Suite 1453, San Francisco, CA 94104 Tel: +1 415 403 9027
MUNICH: Nymphenburger Strasse 14, D-80335 München Tel: +49 89 549057-13
PARIS: Moulin de Malicorne, Sainte Maure de Touraine, Paris 37800 Tel: +33 2 47 35 72 15
TOKYO: Chiyoda Building, 2-37 Ichigayatamachi, Shinjuku-ku, Tokyo 162-0843 Tel: +81 3 3267 8751
INDIA: Nature Publishing Group, 3A, 4th Floor, DLF Corporate Park, Gurgaon 122002 Tel: +91 124 2881053/54

Nature® (ISSN 0028-0836) is published by Nature Publishing Group, a division of Macmillan Publishers Ltd (The Macmillan Building, 4 Crinan Street, London N19XW). Registered as a newspaper at the British Post Office. North and South American orders to subscriptions@natureny.com (Tel: +1 866 363 7860). Other orders to subscriptions@nature.com (Tel. +44 (0)1256 329242; Fax +44 (0)1256 812358). Authorization to photocopy material for internal or personal use, or internal or personal use of specific clients, is granted by *Nature* to libraries and others registered with the Copyright Clearance Center (CCC) Transactional Reporting Service, provided the relevant copyright fee is paid direct to CCC, 222 Rosewood Drive, Danvers MA 01923, USA. Identification code for *Nature*: 0028-0836/03. CPC PUB AGREEMENT #40032744. In the US, *Nature* (ISSN 0028-0836) is published weekly on Thursday, except the last week in December by Nature Publishing Group, 75 Varick St, 9th Fl, New York NY 10013-1917, USA. US Periodicals postage paid at New York NY, and additional mailing post offices. US POSTMASTER: send address changes to *Nature*, Subscription Dept, 75 Varick St, 9th Fl, New York NY 10013-1917, USA. Published in Japan by NPG Nature Asia-Pacific, Chiyoda Building, 2-37 Ichigayatamachi, Shinjuku-ku, Tokyo 162-0843, Japan.

© 2013 Macmillan Publishers Limited. All rights reserved.

nature

A key thread in your scientific career

Whether in print, online or mobile, *Nature* is always at the forefront of science and is your forum to read, watch, listen and engage with key research, news and opinion.

How will you access the next landmark discovery?

➲ NATURE.COM/NATURE

nature publishing group npg

Scientific Writing and Publishing Workshops

Nanjing Medical University

and

Macmillan Science Communication (MSC),
an exclusive partner of Nature Publishing Group,

are holding a Scientific Writing and Publishing Workshop in Nanjing
23rd February 2013

- The MSC intensive one-day training course will train 60 researchers with minimal experience of publishing research papers, how to prepare and publish high quality research papers in international, peer-reviewed journals.
- Training will be led by *Nature Genetics* Chief Editor, Dr. Myles Axton, and former *Nature* and *Nature Genetics* Editor, Dr. Wayne Peng, who bring their expertise and editorial insight from working at a prestigious scientific publishing house.
- The workshop will cover the writing and publishing process from start to finish, using short lectures punctuated with practical exercises and group discussions that focus on the publishing issues that researchers face today.
- On the 22nd February, MSC will provide independent feedback on pre-selected, unpublished papers written by workshop attendees in face to face consultations.

This workshop is brought to you by Macmillan Science Communication, exclusive partner of Nature Publishing Group, publisher of *Nature*.

Nanjing Medical University | **www.njmu.edu.cn**
MSC workshops | **www.macmillanscience.com/training**

nature publishing group npg

THIS WEEK

EDITORIALS

BEHAVIOUR Return to the wild offers clues to animal instinct **p.272**

GERM-FREE Parasitic wasp disinfects insides of cockroach host **p.275**

WORLD VIEW China can inspire with a new vision of carbon trading **p.273**

Troubling thoughts

A sustained commitment to mental-health treatment for Fukushima evacuees could also help survivors of future disasters.

Natural disasters create bold images: winds blow, waves crash ashore, buildings burn. Yet some of the most important long-term consequences are invisible. Survivors can be gripped by anxiety, depression and post-traumatic stress disorder (PTSD) for years after a disaster — especially if they are unable to return home and rebuild.

This is exactly the situation faced by thousands of residents of eastern Fukushima prefecture in Japan, who fled their homes ahead of the meltdown of three reactors at the Fukushima Daiichi nuclear plant in March 2011. As described on page 290, many evacuees are now anxious and depressed about their long-term prospects. Researchers have found evidence that they suffer from elevated levels of anxiety and PTSD-related symptoms, and there are fears that substance abuse and other problems may be on the rise.

Mental health is a major component of the Fukushima Health Management Survey, the government effort to monitor survivors for adverse health effects from the accident. The survey includes a small but competent team of mental-health professionals who are chronicling mental-health and lifestyle issues. Last year, they heard from more than 90,000 evacuees who answered a brief questionnaire about their mental state. The results were not encouraging: levels of PTSD-like symptoms nearly a year after the accident were similar to those of workers who had responded to the 2001 attacks on the World Trade Center in New York.

The health survey's goal is not simply to record the effects of the accident, but also to support the evacuees. Counsellors targeted nearly 5,000 individuals for follow-up phone calls to discuss any mental-health problems. Unfortunately, a phone call may not be enough. The survey itself had less than a 50% response rate, and those who were called for follow-ups tended to stay on the line for just a few minutes — unwilling or unable to talk in detail about their problems.

The researchers involved in the health survey would like to do more. They want to conduct face-to-face interviews and set up counselling centres. But money is tight and the survey is already over budget. It currently spends about twice what the government has allocated, and budgetary wrangling between the prefectural and central governments could lead to further cuts. In this precarious financial environment, the team has found itself unable to hire long-term staff, or even print pamphlets about mental-health issues for the evacuees.

That is an unfortunate state of affairs because Fukushima presents a good opportunity to learn about the best ways to treat the mental-health problems of disaster victims. Although researchers have chronicled mental-health impacts from diverse crises ranging from 2005's Hurricane Katrina and the 2010 earthquake in Haiti to the ongoing Israeli–Palestinian conflict, the literature on how to actually treat these problems is fairly thin. Fukushima's evacuees are similar to survivors of other events in many ways. Nearly all suffer from sub-clinical symptoms that are difficult to diagnose. The size of the afflicted population makes one-on-one therapy impractical. And the situation is complicated by a deep mistrust of the authorities, whom many evacuees blame for their present woes. But unlike many disaster zones, Fukushima is a developed region with a well-educated, well-documented and contactable population. Much could be learned by studying and treating these evacuees in the long term.

And the evacuees will need that help. Unlike the survivors of many other disasters (including those who were affected by the tsunami that sparked the nuclear meltdown), Fukushima's nuclear evacuees live in fear of radiation exposure. They worry that they or their children may fall ill from the accident, and their anxiety could grow as the years pass. Nearly half of all Japanese people will develop cancer at some point in their lives, and evacuees will wonder whether that cancer is connected to the accident. Studies of mothers who were evacuated from the 1986 nuclear disaster in Chernobyl, Ukraine, show that this anxiety for their children can last for decades, and may never entirely go away.

> ***"Although researchers have chronicled mental-health impacts from diverse crises, the literature on how to treat these problems is thin."***

Given all this, the Fukushima Health Management Survey deserves continued and even increased support from the government. Survey scientists should also seek lasting collaborations with researchers outside Japan so that the lessons learned there, so painfully earned, can be shared with the world. ■

Knowledge trades

Institutions must carefully evaluate their researchers' relationships with Wall Street.

Few researchers would care to put a price on their professional reputation. Yet when neurologist Sidney Gilman decided to divulge confidential clinical-trial data to a hedge-fund manager, he did exactly that. Court documents show that Gilman earned more than US$100,000 for his illegal tips about the failure of an experimental drug for Alzheimer's disease; the hedge fund made $276 million in direct gains and avoided losses. After he was caught, Gilman lost not only the cash but also his career, retiring from his position at the University of Michigan Medical School in Ann Arbor.

In recent conversations with a *Nature* reporter, several academics converged in their assessment of Gilman's case,which became the biggest insider-trading case in US history: "stupid".

Nevertheless, many of these same academics continue to serve as consultants for the investment industry, where knowledge is money and confidential knowledge is the most precious currency of all. Like Gilman, they participate in 'expert network' firms that bring together academic specialists and clients who seek technical information (see page 280). Such firms do not work exclusively for the investment industry, but hedge funds make up a sizeable part of their business. Expert networks make it easier for researchers to dabble at advising Wall Street, often for clients who do not disclose the companies for which they work.

Undoubtedly, the vast majority of consultations for financial firms do not result in illegal exchanges. Gilman and his alleged hedge-fund co-conspirator were clearly mindful of their transgressions and went to some lengths to evade the barriers to insider trading erected by the expert network that united them in the first place. Experts contacted by *Nature* were confident that they had never divulged confidential information. They were less confident about others, noting that less experienced or over-eager colleagues might be prone to a slip of the tongue here and there.

A small slip can move markets. Telltale cues from body language (a shift in a chair) and tone of voice (a hesitation, a cough) can speak volumes, as can the unguarded answer to one acute question slipped into an otherwise innocuous conversation. Indeed, journalists might find the techniques used by hedge-fund managers to dig out confidential information uncomfortably familiar. Many academics are trained to handle the press; few receive education in how to deal with the financial industry.

The cavalier attitude towards this work is disturbing. For busy physicians, recruitment letters from expert networks are part of a steady flow of surveys and consulting requests that clog their inboxes. Some of the researchers interviewed by *Nature* could not remember which expert networking firms they consult for. Consultations were often viewed as an easy way to pick up a little extra cash when time permits, and, if lucky, perhaps have an engaging conversation along the way.

This casual approach extends to institutions, many of which are well versed in negotiating the rocky road of conflicts of interest raised by consultations for the drug or medical-device industries, but have not explored the issues raised when advising hedge funds. Universities are already wrestling with the mounting requirements from federal funders for the reporting of potential conflicts of interest, and are unlikely to welcome yet another category of extracurricular activities to monitor.

All involved must take the relationship between researchers and Wall Street more seriously. Institutions should discuss the risks involved and, when warranted, take a proactive stance, perhaps using the Cleveland Clinic in Ohio as a guide. Since 2005, the clinic has instituted a special level of legal review for relationships between faculty members and the investment industry. Physicians who embark on these relationships are given special educational material — which might soon include the newspaper accounts of the Gilman case.

"Avoidance may be the best strategy to prevent accidental leaks."

As US regulators cracked down on insider trading, some researchers cut their ties with the financial industry and expert networks for fear of being tainted by association. Those who still consult for Wall Street often say that they do so to help guide investment in their field. That aim is laudable, but it cannot be used to justify consultation on topics that could overlap with privileged information. For example, a researcher who works on a clinical trial for an experimental diabetes drug should think twice before consulting with a hedge fund about diabetes drugs.

Avoidance may be the best strategy to prevent accidental leaks. The Gerson Lehrman Group, the expert network based in New York that employed Gilman, has policies in place to protect consultants from the consequences of unintended disclosure. Two stand out as potentially the most powerful: an academic can refuse a consultation if the subject matter might tread near confidential information; and he or she can abort a consultation — and still get paid — if a client presses for insider knowledge. So the solution is simple: when asked to disclose confidential information, hang up the phone. ■

Natural history

Age-old field methods can tell us more about animal behaviour than can laboratory models.

In the autumn of 1927, the biologist Francis Sumner spent two months in Florida and Alabama trapping wild oldfield mice for studies of skin pigmentation. With the advice of local farmers, Sumner managed to shoo hundreds of rodents out of their burrows.

Sumner also took the opportunity to document the complexity of the vacated mouseholes, detailed in an article published two years later (Sumner, F. B. & Karol, J. J. *J. Mammal.* **10,** 213–215; 1929). The creatures' burrows included a long entrance tunnel leading to a nest and, in case of an invading snake, a secondary escape tunnel that didn't quite reach the surface, a metre or so from the entrance. Other ethologists have since characterized the humble mouse burrow, the structure of which is seen as a model of complex animal behaviour.

Fast-forward almost a century, and a team led by Hopi Hoekstra of Harvard University in Cambridge, Massachusetts, has elegantly unpicked the genetic basis of this behaviour using a cross-breeding design and cutting-edge genotyping methods. Such work, published in this issue (see pages 284 and 402), should appeal to more than just mouse fanciers. Like few papers before, the work shows how long-forgotten field observations, evolutionary theory and molecular genetics can all be brought to bear on a single question.

We have learned much about the physiology of behaviour from model organisms such as laboratory mice — for example, the discovery of genes that determine circadian rhythms, which revealed important mechanisms underlying behaviours such as sleep. But decades of selection for convenient traits such as docility have made laboratory models less than ideal for studying the evolution of complex behaviours. They tell biologists little about the vast behavioural differences that can exist between closely related animals, probably as a result of natural and sexual selection.

Scientists interested in probing the behaviour of wild animals can follow Hoekstra's lead and pick animals and behaviours with a rich history of observation and striking differences between close relatives. Decades-old observations of ant behaviour, including those by Edward O. Wilson, culminated in the discovery, published online in *Nature* today, that a social chromosome explains why some red imported fire ant colonies have one queen, whereas others accept multiple queens.

There are, of course, risks to tackling behavioural genomics in wild animals. The ultimate proof of any gene's role in a specific behaviour involves knocking in or out the gene to remove or endow that behaviour. Such experiments are a challenge even in model organisms, and so far few precedents have been set in non-model species.

Model organisms, imperfect as they are when it comes to studying some behaviours, have focused attention on a handful of organisms. If every interesting animal becomes fair game, there is a risk that behavioural genetics will be fragmented. "If everyone does it in their own species, it will not be a very productive type of enterprise," says Laurent Keller, a geneticist at the University of Lausanne, Switzerland, who led the ant research. He suggests that scientists converge on a set of wild animals in which to intensely study behaviour. If the latest research is any indication, such animals will be no strangers to historians of biology. ■

➲ NATURE.COM
To comment online, click on Editorials at:
go.nature.com/xhunqv

WORLD VIEW *A personal take on events*

China has the capacity to lead in carbon trading

Pilot schemes launched this year could be the start of a world-class system — if the country can solve its data-gathering problems, says **Qiang Wang**.

The value of the world's carbon market fell for the first time last year. More than one-third was wiped from the price of carbon credits in a plunge that reflects the continued global economic crisis and uncertainty over the future of emissions-trading schemes.

Is it a good time to buy carbon credits? Perhaps not yet, but some shoots of recovery are visible, not least in my own country of China. Largely unnoticed in the West, Chinese carbon trading is getting up and running. In just two years, officials have designed and started to implement seven trading trials that cover around one-third of China's gross domestic product and one-fifth of its energy use. If successful, the schemes will show that emissions trading could be a powerful way for China to control its greenhouse-gas emissions.

The first trades in one of these schemes took place in September 2012, when four cement-manufacturing companies in the southern, industrial region of Guangdong province invested several million dollars each in carbon-pollution permits, which they will need to expand operations. The Guangdong scheme is expected to cover more than 800 companies that each emit more than 20,000 tonnes of carbon dioxide a year across nine industries, including the energy-intensive steel and power sectors. These firms account for more than 40% of the power used in the province. The Guangdong carbon market alone will regulate some 277 million tonnes of CO_2 emissions by 2015, almost equal to Ukraine's total annual CO_2 emissions.

CHINA'S POLITICAL SYSTEM COULD LET A **CARBON MARKET** GROW FASTER THAN **ANYWHERE ELSE.**

China plans to open six further regional emissions-trading schemes this year, in the province of Hubei and in the municipalities of Beijing, Tianjin, Shanghai, Chongqing and Shenzhen. It plans to expand and link them until they form a nationwide scheme by the end of the decade; that would then link to international markets.

China is not alone. Australia and South Korea are scheduled to open their own carbon markets in 2015, and California launched its first carbon-allowances auction in November 2012.

China may not seem a natural home for carbon trading. With heavy government intervention, significant state ownership of enterprises and a culture of distrust in business, the country remains far from a true market economy. Until now, China's experience of carbon trading has been almost exclusively under the Clean Development Mechanism, the international carbon-trading scheme set up under the Kyoto Protocol, to which it is the leading supplier of credits. Only a few symbolic transactions of carbon offsets have been made since the Beijing, Tianjin and Shanghai voluntary emissions-trading exchanges opened in 2008.

➲ NATURE.COM
Discuss this article online at:
go.nature.com/c93naw

However, China's political system could let a carbon market grow faster than anywhere else. Once Chinese leaders have accepted a concept, opposition is steamrollered and changes are implemented much more quickly and broadly than is possible in societies in which policy-making is based on a balance of the interests from different stakeholders. Stable policies are crucial to the success of emissions trading, which — unlike most markets — grows from the top down.

Challenges lie ahead for emissions trading in China. The country needs to develop and enforce proper legislation and regulations to measure, report and verify carbon emissions from industrial sites. It needs to build an effective and accountable framework to oversee the reporting and trading of carbon credits. Most urgently, China needs to look at how it collects and analyses data on carbon emissions.

The credibility of China's statistics on energy use and carbon emissions has long been questioned, partly because numbers calculated using top-down and bottom-up statistics do not match. The discrepancies can be very large. For example, last year, scientists compared Chinese CO_2 emissions as calculated using top-down energy data from the National Bureau of Statistics of China with those calculated using bottom-up data from the 30 provincial statistics bureaus. For 2010, the bottom-up figure was larger by an amount equivalent to the total annual CO_2 emissions of Japan, about 5% of the global total (D. Guan *et al. Nature Clim. Change* **2,** 672–675; 2012). Without accurate numbers, the first deal of the Guangdong trading scheme was based on expected future carbon emissions, rather than historical data.

China should set up a reliable system for gathering carbon data as soon as possible, and build a comprehensive database of carbon emissions. Standard international methods for carbon trading to reduce emissions demonstrate how to set up a baseline, show that any emissions change is genuine and calculate the reductions.

Reliable data-gathering will require effort from the scientific community at home and abroad, to improve statistical methodology. It will also need political action. The organizations that collate and publish carbon statistics should be made independent of possible government interference. And China needs specific laws to ensure honest reporting and transparency, and to punish those who make fraudulent or misleading claims about their carbon emissions.

If China can build a workable and credible emissions-trading scheme, it will benefit not just one country, but the entire world. ■

Qiang Wang *is director of the Western Research Center for Energy and Eco-Environmental Policy at the Xinjiang Institute of Ecology and Geography of the Chinese Academy of Sciences in Urumqi.*
e-mail: wang_q@ms.xjb.ac.cn

RESEARCH HIGHLIGHTS

Selections from the scientific literature

BIOCHEMISTRY

Machine mimics ribosome

The cell's protein-building complex has inspired the development of a molecular machine that links amino acids together.

David Leigh at the University of Manchester, UK, and his colleagues based their 'nanomachine' on the ribosome, a complex of proteins and RNA that translates the genetic code into proteins. But, compared with the ribosome, their machine is primitive and slow, and can assemble only very short protein chains called peptides. The nanomachine is based on a rotaxane, which is a large molecular ring threaded onto a molecule that acts as an axle. In the experiment, the axle is lined with three amino acids, and a chain of three amino acids hangs from the outer edge of the ring. When the machine is heated, an amino acid from the axle is transferred to the end of the hanging chain. The ring can then move along the axle, repeating the reaction.

Running about 10^{18} molecular machines at once produces tens of milligrams of peptide.

Science **339, 189–193 (2013)**

For a longer story on this research, see go.nature.com/p3rtai

ED RESCHKE/PETER ARNOLD/GETTY

ECOLOGY

Cicadas emerge when predators decline

Periodical cicadas, which spend most of their lives underground and occasionally emerge en masse, may time their appearances with periodic dips in predator numbers.

Why several species of the insect (*Magicicada* spp; pictured) surge to the surface of North American forests after 13 or 17 years underground has been a mystery. Walt Koenig at Cornell University in Ithaca, New York, and Andrew Liebhold of the US Department of Agriculture Forest Service in Morgantown, West Virginia, analysed data on the population sizes of 15 bird species between 1966 and 2010 to estimate the predation pressure on the insects. They conclude that the cicadas somehow set their predators' populations into cycles that reduce bird numbers during their next foray above ground.

This could be explained by the insects' long-lasting effect on forest ecology influencing factors such as nutrient levels and tree growth for several years after their emergence, the authors suggest.

Am. Nat. 181, **145–149 (2013)**

HUMAN EVOLUTION

Migration from India to Australia

A genetic analysis indicates that humans migrated from India to Australia around 4,000 years ago. This contradicts the prevailing view that, after its initial colonization, Australia had little contact with the rest of the world before the late eighteenth century.

Irina Pugach at the Max Planck Institute for Evolutionary Anthropology in Leipzig, Germany, and her colleagues analysed genome-wide genetic variation from 344 people, including Australian Aboriginals and individuals from Africa, India and island Southeast Asia. The authors found evidence of a migration out of Africa into Australia at least 36,000 years ago, and one from India about 4,230 years ago.

The Australian archaeological record shows rapid changes in tool technology around 4,000 years ago, leading the authors to suggest that the migrants could have brought in the new technology.

Proc. Natl Acad. Sci. USA **http://dx.doi.org/10.1073/pnas.1211927110 (2013)**

NEUROSCIENCE

Hearing restored with new hair cells

An experimental drug can restore hearing in deaf mice by regenerating sound-sensitive cells in the inner ear.

Excessive noise permanently damages hair cells that conduct sound to the brain and are found in an inner-ear structure called the cochlea. A team led by Albert Edge at Harvard Medical School in Boston, Massachusetts, has identified a molecule that can convert other cochlear cells into hair cells. The compound, named LY411575, blocks a biochemical pathway called Notch, which normally prevents supporting cells in the cochlea from developing into hair cells. When applied to the ears of mice deafened by loud sounds, the molecule triggered the generation of new hair cells. By three months after this treatment, the rodents' hearing had been partly restored.

LY411575 would probably need to be delivered locally to the ear in humans to avoid side effects, the authors say.

Neuron 77, **58–69 (2013)**

ANIMAL BEHAVIOUR

Mothers call for parenting help

Male laboratory mice do not spontaneously care for their offspring, but they can exhibit some paternal behaviour in response to signals from their mates.

Haruhiro Higashida at Kanazawa University in Japan and his colleagues found that when female mice were separated from their young, they could encourage their male partners to adopt parental behaviour — such as carrying wandering pups back to the nest (**pictured**) — using high-pitched vocalizations and pheromones.

About 60% of the males tested retrieved pups after hearing a recording of 38-kilohertz distress calls of females removed from their pups. The team found similar results when males were placed in cages that had previously held females separated from their young. Paternal behaviour disappeared when the researchers simultaneously blocked hearing and olfaction in the males.

Nature Commun. 4, **1346 (2013)**

ARCO IMAGES GMBH/ALAMY

MATERIALS

Spinning strong, conductive carbon

Fibres spun from carbon nanotubes can be as electrically conductive as metal wires, yet as strong as conventional carbon fibres.

Individual carbon nanotubes are strong, stiff and exceptionally conductive — but spinning them into larger filaments adds defects, impurities and misalignments that compromise the fibres' physical properties. Matteo Pasquali at Rice University in Houston, Texas, and his colleagues have improved the spinning process by adapting techniques used to make industrial fibres such as Kevlar. The authors extruded filaments from nanotubes dissolved in an acid and wound them into fibres of dense, well-aligned tubes that are stronger than copper and almost as conductive.

The fibres could be manufactured on a large scale and have a variety of applications in electronics, the authors say.

Science 339, **182–186 (2013)**

CELL FATE

Impede a protein to reprogram

Blockade of just one protein is sufficient to change skin cells into neurons.

Several research groups have already converted one cell type to another by adding, for example, specific genes. Now researchers led by Xiang-Dong Fu at the University of California, San Diego, have found another approach. They used small RNA molecules to repress an RNA-binding protein called PTB. Decreased levels of this protein led to the activation of genes that can convert skin cells to neurons. PTB influences the stability of RNA molecules and coordinates molecules called microRNAs that regulate gene expression.

The approach is a convenient way to reprogram other cell types to neurons, and provides a glimpse of the complex mechanisms involved in the control of gene expression.

Cell **http://dx.doi.org/10.1016/j.cell.2012.11.045 (2013)**

COMMUNITY CHOICE

The most viewed papers in science

CLIMATE SCIENCE

Warming matches predictions

HIGHLY READ on iopscience.iop.org in December

Temperature data collected since 1990 are in good agreement with model-based estimates of global warming made by the Intergovernmental Panel on Climate Change (IPCC).

Stefan Rahmstorf of the Potsdam Institute for Climate Impact Research in Germany and his colleagues removed short-term temperature fluctuations or 'noise', which result from events such as volcanic eruptions, from a 1980–2011 time-series of annual global temperatures. The adjusted warming trend of 0.16 °C per decade closely matches the projections made by the IPCC in its reports in 2001 and 2007.

The global sea level, however, has since 1993 been rising 60% faster than anticipated. Future sea-level rise could exceed even the highest value — some 60 centimetres by the end of the century — projected by the IPCC in its 2007 report, the team cautions.

Environ. Res. Lett. 7, **044035 (2012)**

GUDRUN HERZNER/PROC. NATL ACAD. SCI.

ANIMAL BEHAVIOUR

Wasp parasites keep hosts clean

By peering into the bodies of American cockroaches, researchers have found that a parasitic larva disinfects its host with antibacterial secretions.

Larvae of the emerald cockroach wasp (*Ampulex compressa*; **pictured**) feed on the innards of the American cockroach (*Periplaneta americana*), but also have to contend with bacteria that live in their host's tissues. Gudrun Herzner at the University of Regensburg in Germany and her colleagues installed transparent panels into the sides of parasitized roaches and observed that the wasp larvae secrete large amounts of a clear liquid from their mouthparts onto their hosts' tissues.

Analysis of the liquid revealed the presence of the chemicals mellein and micromolide. These substances slow the growth of certain microorganisms, including the bacterium *Serratia marcescens*, which can kill insect larvae and was also isolated from the cockroaches.

Proc. Natl Acad. Sci. USA **http://dx.doi.org/10.1073/pnas.1213384110 (2013)**

For a longer story on this research, see go.nature.com/vxttfo

➲ NATURE.COM
For the latest research published by *Nature* visit:
www.nature.com/latestresearch

SEVEN DAYS *The news in brief*

MARK RALSTON/AFP/GETTY

Record smog blankets Beijing

A year after Beijing's municipal government said that it would report more realistic assessments of the city's air quality, pollution in China's capital soared to unprecedented heights. City officials last week asked construction sites and industries to limit their activities and told residents to stay indoors, as concentrations of fine particles (those less than 2.5 micrometres across) rose to record levels many times above what the World Health Organization considers 'hazardous'. With smog blanketing cities across northern China, state media urged the government to take action on air pollution.

RESEARCH

Antarctic lakes

A team of US-led Antarctic explorers was preparing to drill through to the subglacial Lake Whillans as *Nature* went to press this week. The pristine lake is buried beneath some 800 metres of ice at the margin of the West Antarctic Ice Sheet. Meanwhile, Russian scientists who last February reached Antarctica's largest subglacial lake, Lake Vostok, announced on 10 January that they had recovered further samples of lake water, which had moved up the borehole and refrozen onto the drill bit. See go.nature.com/ut2hxw for more.

US climate audit

A US climate-assessment advisory committee released its draft report on 11 January, documenting a range of global-warming impacts across the country and making projections. Issued under the auspices of the inter-agency Global Change Research Program, the document is available for public comment until 12 April, and will serve as the basis for the United States' third national climate assessment, which is expected to be released late this year or early in 2014. See go.nature.com/k9p4lm for more.

EVENTS

Bush-fire damage

A fierce bush fire threatened Australia's Siding Spring Observatory in New South Wales last week, but initial assessments indicated that no telescopes were seriously damaged, the Australian National University (which runs the site) said on 14 January. Some buildings were severely affected at the observatory, which will be closed for two weeks. The fire was one of many across Australia, where long-standing temperature records were broken last week during a summer heatwave.

Warmest year

The contiguous United States chalked up its warmest year on record last year, blowing past the previous record set in 1998. The average temperature, 12.9 °C, was 1.8 °C above the average for the twentieth century and 0.56 °C above the previous record. Global data are not yet available, but the latest analysis by the US National Oceanic and Atmospheric Administration suggests that, in 2012, January to November was the eighth warmest period on record.

POLICY

GM transparency

The European Food Safety Authority announced on 14 January that it was starting a broad transparency initiative, designed to make data from its risk assessments more widely available. The agency also released the first batch of data: a file containing documents and data submitted by biotechnology giant Monsanto for the 2003 authorization of its genetically modified maize (corn) NK603. See go.nature.com/e7sgdd for more.

Primates grounded

United Airlines, the world's largest carrier, will no longer ship non-human primates to research labs. The airline clarified its policy on 8 January, following pressure from activists at People for the Ethical Treatment of Animals. With the adoption of similar rules by Air Canada last month, only four major airlines now say they will fly research primates: Air France, China Eastern Airlines, Philippine Airlines and Vietnam Airlines. See go.nature.com/agkgbl for more.

Japanese windfall

Japanese scientists may benefit from the ¥10.3-trillion (US$120-billion) supplementary budget proposed by the Liberal Democratic Party, which regained power in December and seems set on reversing some of the austerity that has characterized Japan's leadership over the past few years. Of

¥570 billion requested by the science and education ministry, universities will receive ¥180 billion for technology transfer, and research on induced pluripotent stem cells could receive ¥20 billion. The budget is expected to be passed next month. See go.nature.com/7lhmxg for more.

California increase

California's public universities, which have suffered years of funding cuts, were awarded a US$500-million boost in the state's proposed 2013–14 budget, announced on 10 January. The generosity comes after voters last year approved a temporary tax increase to stave off deep cuts to public education. Despite the rise, general state funds stand more than $1.2 billion lower than in 2007–08, a shortfall that universities have made up for by increasing student tuition fees (see *Nature* **480,** 164; 2011). Both Democrats and Republicans expressed support for the budget, which the state legislature must approve before it comes into effect in July.

Red–Dead link

A proposal to construct a 180-kilometre-long conduit from the Red Sea to the Dead Sea is economically and environmentally feasible, according to draft studies posted online last week. Israel, Jordan and the Palestinian Authority had asked the World Bank to look at the plan in 2005, with the idea that channelling water downhill to restore the salty and rapidly shrinking Dead Sea would also generate hydroelectricity to run a desalination plant (see *Nature* **464,** 1118–1120; 2010). The World Bank recommended a pipeline as the best option, at a cost of around US$10 billion. See go.nature.com/if8uyu for more.

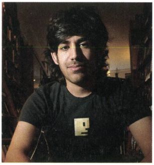
US energy hub

The US Department of Energy has launched its fifth Energy Innovation Hub, a research centre focusing on rare-earth minerals and materials that are crucial to the energy system, including advanced batteries, wind turbines and solar panels. Based at the department's Ames Laboratory in Iowa, the Critical Materials Institute will receive up to US$120 million over five years. The institute, announced on 9 January, follows earlier hubs focusing on advanced nuclear reactors, energy storage, energy-efficient buildings and the development of fuels from sunlight. The hub model was the brainchild of energy secretary Steven Chu.

PEOPLE

Open-access tribute

A campaign on Twitter is calling for researchers to post their papers openly online, at hashtag #pdftribute, in commemoration of Aaron Swartz, a 26-year-old programmer and Internet activist who committed suicide on 11 January. Swartz (**pictured**), who pushed for open access to court records, scholarly material and other web information, was facing criminal charges for downloading some 4 million articles from the not-for-profit scholarly archive JSTOR. Academics have joined a wave of indignation at the charges, which many argue contributed to his suicide. See go.nature.com/5wmeld for more.



NOAH BERGER/REUTERS

BUSINESS

Energy spending

Global investment in clean energy fell by 11% to US$268.7 billion in 2012, according to figures from analysts Bloomberg New Energy Finance on 14 January. Around $30 billion of investment went to research and development, with the rest mainly financing energy projects. China's spending rose 20% to $67.7 billion, mostly owing to greater investment in solar energy. Policy changes such as expired subsidy programmes led to lower investment in the United States (down 32% to $44.2 billion), Italy (a drop of 51% to $14.7 billion) and Spain (a 68% reduction to $3 billion).

Arctic drilling

US secretary of the interior Ken Salazar announced a federal review of the offshore drilling programme in the Arctic on 8 January, focusing on a series of problems encountered by oil-and-gas giant Shell. Late last month, Shell lost control of a drilling rig, which ran aground in the Gulf of Alaska; the company briefly lost control of another rig in July. The review will focus on these issues as well as on challenges that the company has faced in implementing its oil-spill-preparedness plan. A report is scheduled for completion in early March.

COMING UP

20–25 JANUARY

In Tromsø, Norway, politicians and researchers discuss the geopolitics and changing ecosystems of the Arctic, in the 'Arctic Frontiers' conference.
go.nature.com/kjorrb

24 JANUARY

The Natural History Museum in London launches its year-long celebration of the work of Alfred Russel Wallace — the co-discoverer of evolution by natural selection. This year is the centenary of his death.
go.nature.com/dxbt9r

➲ NATURE.COM

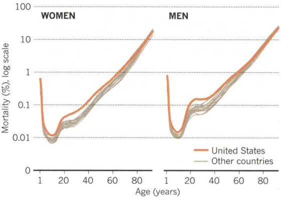
For daily news updates see:
www.nature.com/news

TREND WATCH

Human health and life expectancy are poorer in the United States than in other high-income nations, according to a 9 January report by the National Research Council in Washington DC. Compared with 16 peer countries, including France, Japan and the United Kingdom, the United States has some of the highest rates of infant mortality and obesity. The death rate in the United States is higher than average for most ages, with the largest disparities at younger ages (see chart).

POOR HEALTH IN THE UNITED STATES

Compared with 16 other rich countries, the United States had the highest death rates in most young age categories in 2005–09.



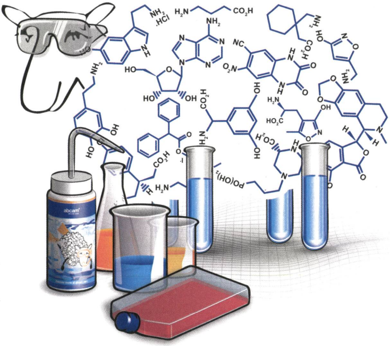
SOURCE: HUMAN MORTALITY DATABASE

Quality products
at exceptional
value

abcamBiochemicals®

Take your research further by exploring our expanding range:

- **Enzyme inhibitors**
  Akt, MAPK, Caspases
- **Signaling tools**
  Apoptosis, Ca^{2+} signaling,
  Endocytosis & exocytosis
- **Receptor agonists and antagonists**
  Glutamate, GABA, Serotonin
- **Ion channel modulators**
  Ca^{2+}, K^+, Na^+



What to expect when you order Abcam Biochemicals:

✓ High purity and exceptional quality guaranteed

✓ Competitively priced

✓ Excellent supporting data

✓ Friendly and knowledgeable customer service

Your quality guarantee • Your quality guarantee •

Discover more at abcam.com/biochemicals

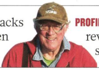
NEWS IN FOCUS

ETHICS Ties between scientists and financial markets are questioned **p.280**

TECHNOLOGY Quantum dots find a future — in television screens **p.283**

GENETICS Sequencing tracks roots of behaviour seen in nature **p.284**

PROFILE Fathering a revolutionary idea, and a scientific dynasty **p.286**



TOMAS VAN HOUTRYVE/VII/CORBIS

New treatments could help to safeguard the health of infants like this baby, born to an HIV-positive mother in Uganda.

AFRICA

HIV trial under scrutiny

Critics say that antibody therapy is too expensive for its African target population.

BY MEREDITH WADMAN

A potential breakthrough in the quest to prevent HIV and AIDS has collided with sensitivities about testing expensive drugs in poor parts of the world. Three years ago, scientists at the US National Institute of Allergy and Infectious Diseases (NIAID) in Bethesda, Maryland, isolated a pair of potent monoclonal antibodies that neutralized more than 90% of the HIV strains that infect people (X. Wu *et al. Science* **329,** 856–861; 2010). Now, one of those antibodies, VRC01 (named after the institute's Vaccine Research Center, VRC), is being readied for a clinical trial in Africa.

But the proposed trial — in 3,000 African infants born to breastfeeding, HIV-infected mothers — is drawing fire from critics. They cite the therapy's steep cost and lack of proven efficacy in adults, and say that an affordable way to prevent mother-to-child HIV transmission already exists. Those points are likely to surface at a meeting in Entebbe, Uganda, on 22–23 January, where attendees will pound out principles for conducting prevention trials in breastfeeding babies born to HIV-positive mothers in poor countries.

One of those invited to the meeting, Arthur Ammann, president of Global Strategies for HIV Prevention in Albany, California, says that the trial would be testing a treatment on infants "that would never be available to them or anyone else in a resource-poor country". But Barney Graham, a senior investigator at the VRC and one of the leaders of the proposed study, asks: "If you have a product you think has a chance of preventing infection, is it ethical not to do that study?"

The impetus for the trial arises from a conundrum: breast milk protects infants from pneumonia and gastroenteritis, major killers in developing countries, but the milk can also transmit HIV. It is estimated that, in sub-Saharan Africa, 40–50% of the roughly 300,000 new infant HIV infections in 2011 were acquired through breastfeeding. But a cocktail of HIV-fighting drugs known as antiretroviral (ARV) therapy can substantially reduce the risk, so the World Health Organization recommends that HIV-positive women breastfeed their infants and that mother and baby both take ARV drugs.

However, ARV therapy is not perfect: infections break through at a rate of 2–5% by the time babies are six months old. The antibodies, which stop the virus from attaching to and entering human cells, offer hope, says ▶

▶ Graham. "Can we get rid of that residual infection by adding an antibody, and would it work well enough to be cost-effective and logistically feasible?" he asks. The trial is designed to answer that question by giving the mother–baby pairs oral ARV therapy while half of the infants also receive monthly injections of VRC01. A control group would get a placebo.

The study will proceed only if VRC01 is approved as an investigational drug by the US Food and Drug Administration, and it would first be safety-tested in adults and infants in the United States. Investigators then hope to move to trial sites in Botswana, Malawi, South Africa, Tanzania, Uganda, Zambia and Zimbabwe. But the critics say that even if the antibody reduces the HIV infection rate, it is unlikely to be widely used in Africa. Whereas ARV therapy is delivered as oral tablets that cost less than US$200 per patient per year in poor and middle-income countries, the antibody must be injected by trained staff. In the United States, monoclonal antibody treatments commonly cost thousands of dollars per year.

Physician and AIDS expert David Ho of the Aaron Diamond AIDS Research Center in New York City is also planning to test an anti-HIV monoclonal antibody, but says that his group decided against running their trial in newborns after working in a poor area in southern China. There, he found that using ARV therapy as well as providing clean water and formula so that women do not have to breastfeed reduced mother-to-child HIV transmission to 1% (Z. Zhou *et al. J. Acquir. Immune Defic. Syndr.* **53** (suppl. 1), S15–S22; 2010). With a transmission rate that low, Ho says, "you are not going to be able to do a meaningful study". Instead, he is planning to conduct his group's prevention study in men who have sex with men.

But Graham says that discouraging breastfeeding, with its many health benefits, is not an acceptable route for reducing infection risk. And Catherine Hankins, deputy director of science at the Amsterdam Institute for Global Health and Development, says that it is premature to worry about cost. She points out that the basic ARV tablets used in Africa dropped in price from more than $10,000 to as little $140 annually in the past decade. "To say up front that something is too expensive, forget it — there are a lot of things we wouldn't have today if people had said that." ■

FINANCE

Insider trading sparks concerns

Universities indulge researchers' ties to finance industry.

BY HEIDI LEDFORD

It is as predictable as a heartbeat: when crucial medical results are at hand, cardiologist Jean-Luc Vachiery knows that the hedge-fund managers will come calling. Vachiery, who tests experimental therapies for a rare condition called pulmonary arterial hypertension at the Free University of Brussels, is privy to confidential information about clinical trials that could be valuable if used — illegally — to guide investments. As such, his conversations with the financial sector can be tense. "They tell you, 'don't say anything that's confidential'," says Vachiery. "And then they ask you for confidential information."

> ***"This is not something that doctors or researchers should be involved in."***

Vachiery, like about 35,000 academics, is a member of the Gerson Lehrman Group, based in New York, one of about 40 'expert network' companies proliferating in the United States and elsewhere that connect clients, often from the financial industry, with experts who can provide technical information. The company recruits heavily from academia and was reluctantly thrust into the limelight last year when one of its most prestigious experts — Sidney Gilman, a neurologist then at the University of Michigan at Ann Arbor — admitted to tipping off a hedge-fund manager about clinical-trial data before they were made public. The result: US$276 million in illicit gains for the hedge fund, the largest insider-trading case the US Securities and Exchange Commission (SEC) has ever handled.

The case is just one in a string of SEC probes launched since 2009 into whether expert networks are trafficking insider information to the financial industry (see 'Trading on expertise'). That investigation has already led to charges against nearly 30 people with connections to expert networks, and is continuing to yield fresh targets, particularly in the health-care industry. It also highlights the tightrope that researchers walk when they consult for the financial industry, for rates that can reach $1,000 per hour. "The only reason anybody wants to talk to you from a financial company is for insider information," says Arthur Caplan, a bioethicist at New York University's Langone Medical Center. "That's the start of the story and that's the end of the story."

Despite the risks, few research institutions have enacted policies to limit staff consultations for the financial industry. Many universities, as well as the US National Institutes of Health, require investigators to disclose consulting income above a given threshold, but few distinguish financial consulting from consulting for drug companies. That's an important distinction, argues Eric Campbell, who studies conflicts of interest at Massachusetts General Hospital in Boston. "This is just a short-term bet for a very select group of people to make a whole lot of money," says Campbell. "This is not something that doctors or researchers should be involved in."

Even so, the networks and the experts they employ maintain that they are doing nothing illegal. Gerson Lehrman, like many other expert networks, has policies to prevent the exchange of confidential information, including a mandatory online training course, and pre-interview questionnaires to prevent experts with insider information from being matched with clients seeking information about that particular field. "All of this makes us a bad place to break the law,"

MORE ONLINE

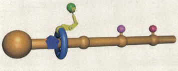
TOP STORY

DANA BERRY/KEPLER/NASA

The most common stars host myriad potentially habitable planets go.nature.com/pxgbbt

MORE NEWS

- Epigenetics drives evolution of invasive species go.nature.com/hwgdis
- Poll finds that experts are split on the meaning of quantum theory go.nature.com/s9vnwe
- Atomic clock based on mass spurs controversy go.nature.com/k5nhfk


MIRIAM WILSON

VIDEO

'Artificial ribosome' nanorobot builds a protein go.nature.com/p3rtai

JOHN MARSHALL MANTEL/THE NEW YORK TIMES/REDUX/EYEVINE

says Alexander Saint-Amand, chief executive of Gerson Lehrman. Experts contacted by *Nature* say that they are often asked for mundane medical information, and are only sometimes asked to divulge secrets. All of them, including Vachiery, say that they never have. "Probably 99% of the use of expert networks has to be considered legitimate," says Michael Mayhew, chief executive of Integrity Research, an investment research firm in New York that tracks the expert network industry.

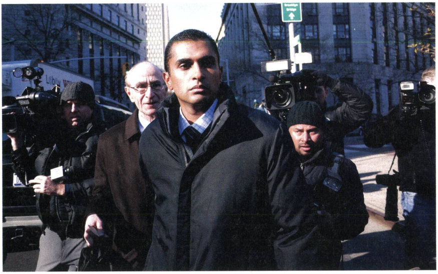
Yet Vachiery says that he is acutely aware of the techniques that sophisticated hedge-fund managers use to dissect his answers. Even as he deflects the questions in face-to-face meetings and teleconferences, he knows that his clients might gather small clues from his body language or his demeanour, patching those clues together with crumbs they've garnered from other experts. "With these bits and pieces they build their story," he says. "They even gauge your nonreaction."

With the risks so high, Mayhew says that he is surprised that hospitals and universities haven't banned participation in expert networks. "If you are afraid that sensitive information might get out, you probably should be proactive."

One of the few medical institutions to monitor financial consulting closely is the Cleveland Clinic in Ohio, which since 2005 has required that all such relationships be reviewed by a lawyer and is now considering an outright ban. Caplan says that professional associations should take a lead role in discouraging members from financial consulting. But Heather Pierce, a director of science policy at the Association of American Medical Colleges in Washington DC, says that she does not know of any professional societies that do.

Marcia Boumil, who oversees conflicts of interest at Tufts University Medical School in Boston, says that the school has no plans to create separate financial consulting disclosure requirements, nor does it they plan to limit researchers' ability to consult. "We're talking about highly respected professionals and we're not really in the business of policing what they do." Nevertheless, Boumil notes that her office might well have noted something odd about Gilman, had he worked at Tufts. Gilman earned $1,000 an hour for his consultations, and made more than $100,000 a year from them — two details that he would have been required to report at Tufts — and that would have raised a red flag.

But David Wazer, a radiation oncologist at Tufts, defends the $1,000 per hour consultations Gerson Lehrman arranges for him. "I'm worth it," he says, adding that, although he often does not know who his clients work for, he has never been asked to give up confidential information. The most he has ever made in one year is $4,000, "not the sort of compensation that would drive me to do anything peculiar", he says.



TRADING ON EXPERTISE

Since 2009, the US Securities and Exchange Commission (SEC) has been cracking down on insider trading by participants in expert networks.

20 NOVEMBER 2012 The SEC charges hedge-fund manager Mathew Martoma (pictured) with insider trading on confidential data gleaned from neurologist Sidney Gilman, whom he met through Gerson Lehrman.
Illicit gains: **US$276 million**.

17 FEBRUARY 2012 The SEC charges Oregon-based expert network firm Broadband Research Corporation and its owner John Kinnucan with trading inside information about 12 technology firms.
Illicit gains: **$110 million**.

3 FEBRUARY 2011 The SEC charges six consultants and employees of Primary Global Research in Mountain View, California, with illegally tipping off hedge funds about technology companies.
Illicit gains: **$6 million**.

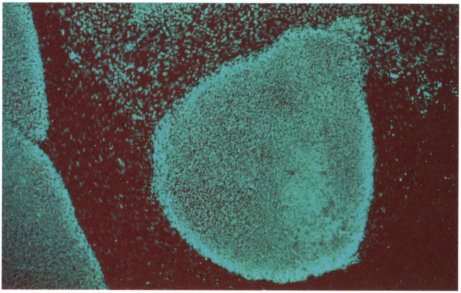
2 NOVEMBER 2010 The SEC charges French hepatologist Yves Benhamou with revealing clinical-trial details to a hedge-fund manager he met through an expert network.
Illicit gains: **$30 million**.

16 OCTOBER 2009 The SEC charges Raj Rajaratnam and his hedge-fund advisory firm Galleon Management in New York of insider trading in the first of a series of cases that eventually implicated people with connections to expert networks.
Illicit gains: **$53 million**.

Instead, Wazer says that his motivation is to advance the development of radiation therapies by influencing investment. "The technologies that move forward are very much a function of where capital flows," he says. "And let's face it, these guys in financial firms determine the flow of capital." Others offer a different explanation: "The pay is good," says Donald Tsai, an oncologist at the University of Pennsylvania's Abramson Cancer Center in Philadelphia, who charges around $300 per hour and says that his clients typically ask him for "kindergarten-grade medical information". "It's fun and easy to do in my free time."

But in the wake of the SEC investigations, some researchers have decided that the stigma attached to expert networks outweighs the pay. Mark Ratain, an oncologist at the University of Chicago in Illinois, severed ties with Gerson Lehrman when the SEC announced its investigation of expert networks. "I didn't want my name associated with this," he says.

And some of those who continue to participate say that they have become wary. Vachiery refuses repeat invitations from those who have prompted him for confidential information. Although he has enjoyed his conversations with some clients, the pressure has made him sour about the whole process. "I'm weighing each word I'm giving to these people and my level of trust is getting lower," he says. "Each time I receive a request now I'm thinking, 'Why am I doing this?'" ■ SEE EDITORIAL P.271



Researchers are keen to compare induced pluripotent stem cells (pictured) with their embryonic cousins.

SILVIA RICCARDI/SPL

STEM CELLS

Court lifts cloud over embryonic stem cells

But research on induced stem cells may be the real winner.

BY MONYA BAKER

The US Supreme Court's decision last week to throw out a lawsuit that would have blocked federal funding of all research on human embryonic stem cells cleared the gloom that has hung over the field for more than three years. Yet the biggest boost from the decision might go not to work on embryonic stem (ES) cells, but to studies of their upstart cousins, induced pluripotent stem (iPS) cells, which are created by 'reprogramming' adult cells into a stem-cell-like state.

At first glance, iPS-cell research needs no help. Researchers flocked to the field soon after a recipe for deriving the cells from adult mouse cells was announced in 2006, partly because this offered a way to skirt the thorny ethical issues raised by extracting cells from human embryos. But the real allure of iPS cells was the promise of genetically matched tissues. Adult cells taken from a patient could be used to create stem cells that would, in turn, generate perfectly matched specialized tissues — replacement neurons, say — for cell therapy. Although the number of published papers from iPS-cell research has not yet caught up with that of ES-cell work (see 'Inducing a juggernaut'), US funding for each approach is now roughly matched at about US$120 million a year.

But, as iPS cells crop up in ever more labs, ES cells — generally cheaper, better behaved and backed by an extra decade's worth of data — promise to have an important supporting role. Ever since iPS cells were described, researchers have been trying to understand just how similar they are to ES cells. iPS cells begin with different patterns of gene expression, and they can also acquire mutations during the reprogramming process, which means that every iPS cell must be thoroughly evaluated before it can be used in any study. "Human ES cells will always be the standard to which other cells will be compared," says Roger Pedersen, who studies how stem cells retain embryo-like states at the University of Cambridge, UK.

Federally supported ES-cell research was shut down in the United States on 23 August 2010, a year after a lawsuit was filed by two opponents of human ES-cell research, and remained frozen for more than two weeks. Many investigators shied away from the field for fear of having to shut down again. The Supreme Court's move has reassured investigators such as Candace Kerr, who studies early development of the brain at the University of Maryland School of Medicine in Baltimore. As a young scientist working towards tenure, she felt particularly vulnerable to the threat of ES-cell funding being stopped. So she switched to iPS cells in 2010, while the lawsuit was working its way through the US court system. With the litigation over, she says she need not "hesitate or fear" adding to her work with experiments using ES cells, which she finds much easier to prompt into neurons than iPS cells. "I am excited and relieved by this decision," she says.

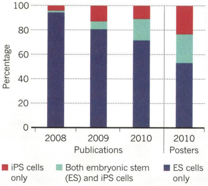
ES cells do, however, have a head start in the clinic. Former heads of the biotech company Geron, based in Menlo Park, California, last week announced an agreement to acquire stem-cell assets including the company's flagship human ES-cell trial, in which precursors of neural support cells grown from human ES cells were injected into people with spinal-cord injuries. In 2011, Geron had shut down the $25-million-a-year effort, after the trial had gone on long enough to show that an ES-cell-based product could be safe. However, it had not shown signs of benefiting patients.

If revived, the spinal-cord trial would double the number of companies sponsoring human clinical trials for ES-cell therapies. At present, the only active trials are those sponsored by Advanced Cell Technology of Santa Monica, California: these are showing early evidence that a product made from human ES cells can help to rebuild the layer of cells that support photoreceptors in the eyes of people with certain types of blindness.

But iPS cells are edging towards the clinic too. Advanced Cell Technology says that it will begin talking to the US Food and Drug Administration this March about the safety studies required to test platelets derived from iPS cells in humans, and Japan is setting up a stem-cell bank of some 75 iPS-cell lines intended for future therapies. Even James Thomson at the University of Wisconsin–Madison, who isolated the first human ES cells in 1998, says that the future belongs to induced stem cells. His lab hasn't bothered to derive a single human ES-cell line since iPS cells were reported. "But we've derived dozens if not hundreds of human induced pluripotent stem cells," he says. ■

INDUCING A JUGGERNAUT

The proportion of stem-cell publications and posters featuring induced pluripotent stem (iPS) cells rose rapidly in the years after their 2006 discovery.



C. T. SCOTT *ET AL. CELL* **145**, 820–826 (2011)

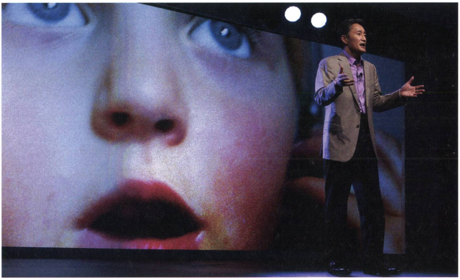
➲ NATURE.COM
For more about the lawsuit, see:
go.nature.com/lypajz

TECHNOLOGY

Quantum dots go on display

Adoption by TV makers could expand the market for light-emitting nanocrystals.

BY KATHERINE BOURZAC


STEVE MARCUS/REUTERS

Sony has announced that it will embed quantum dots in its latest flat-screen televisions.

Live from your living room, in super-saturated colour: it's the quantum-dot TV! Researchers working with nanoscale fluorescent particles called quantum dots have long predicted groundbreaking achievements, such as ultra-efficient light-emitting diodes (LEDs) and solar cells, but the technology has found mainly niche applications. That could change with the announcement last week that QD Vision, based in Lexington, Massachusetts, would supply Sony Corporation of Tokyo with quantum dots for flat-screen televisions that will transmit more richly coloured images than other TVs on the market.

Demand for quantum-dot displays, say industry watchers, could benefit quantum-dot companies, bring down the price of these nanomaterials and boost other applications that have stalled. "Displays are a potential market that could help quantum-dot companies find traction," says Jonathan Melnick, an analyst at Lux Research in Boston, Massachusetts.

Quantum dots are crystals about 10 nanometres in diameter, made from a semiconductor material, commonly cadmium selenide. They are so tiny that their shape and size affect the quantum properties of their electrons, in particular their energy gap — the energy needed to kick electrons into a higher-energy band — which determines the colour of light that the material can emit. Whereas a bulk semiconductor is limited to emitting a single colour of light, researchers can tune the precise colour a quantum dot will absorb and re-emit by tailoring its size.

Discovered in 1981, quantum dots did not find applications until 2002. That was when the Quantum Dot Corporation of Hayward, California, began selling them to cell biologists, who prize them as fluorescent imaging labels for proteins and other biological molecules. As recently as 2010, the biomedical sector was responsible for US$48 million of $67 million in total quantum-dot revenues, according to BCC Research of Wellesley, Massachusetts.

Quantum dots have shown promise for electronics, too — for example in solar cells, in which a mix of quantum dots tuned to absorb different wavelengths of light could capture more of the energy in the solar spectrum. But one hurdle to their exploitation was their temperature sensitivity. Near the backlight of a liquid-crystal display (LCD), for example, temperatures can be around 100 °C. At this temperature, the dots lose efficiency and up to half of their brightness, says QD Vision co-founder and chief technology officer Seth Coe-Sullivan. He says that the company spent a long time tuning the chemistry of its quantum dots to make them stable at higher temperatures.

Moungi Bawendi, a chemist at the Massachusetts Institute of Technology in Cambridge and a co-founder of QD Vision, admits that the company also made some business miscalculations. For its first product, in 2009, it provided Nexxus Lighting of Charlotte, North Carolina, with quantum-dot coatings to convert the harsh glare of LEDs into a warmer glow, to make them more appealing as long-life, low-energy light bulbs. But Bawendi says that LED designs and technology for the light-bulb market evolved too fast for the quantum-dot coatings to keep up. "You design something, and six months later it doesn't work," Bawendi says. "Television technology is more stable."

His optimism will be tested this spring with the company's quantum-dot debut in Sony LCD televisions, to be sold under the Triluminos brand name. The contrast with today's flat screens begins with the light source. Conventional LCDs use a high-intensity blue LED backlight whose glow is converted by a phosphor coating to create a broadband, white light used to make the moving TV images. The new Triluminos televisions instead pair an uncoated blue LED with a thin glass tube filled with quantum dots. Two kinds of quantum dots in the tube absorb some of the blue light from the backlight and re-emit it as pure red and green light. The resulting white light is more intense at the wavelengths of these three specific colours than the white light made by a phosphor-coated LED, so that more colour comes through in the images.

Another quantum-dot company, Nanosys of Palo Alto, California, is providing 3M of St Paul, Minnesota, with material for a similar product. 3M will make a polymer film seeded with quantum dots that does the same job as QD Vision's glass tube. The film is layered between the LCD's stack of light filters, diffusers and polarizers, and similarly converts raw blue light into white light made up of pure colours. Nanosys and 3M announced their partnership in June 2012, but have not yet named any customers.

BCC predicts that, by 2015, optoelectronics, including display components, will make up $310 million of a total $666 million in quantum-dot revenues. Melnick says that these numbers might be overly optimistic, because quantum dots remain expensive. "Even on the low end, they still cost in the hundreds of dollars per gram, and range up to $10,000 per gram," he says. But demand from 3M and Sony could help to bring prices down. Although neither QD Vision nor Nanosys would comment on the volume of material they expect to make this year, or their selling price, both say that they are scaling up their manufacturing volume.

Bawendi is not surprised that it took quantum dots so long to find their footing. "You could argue that 30 years is about the right amount of time from fundamental discovery to applications," he says. ■

GENETICS

Behaviour genes unearthed

Speedy sequencing underpins genetic analysis of burrowing in wild oldfield mice.

BY EWEN CALLAWAY

Even as home experiments go, Hopi Hoekstra's one was peculiar: she built a giant plywood box in her garage in San Diego, California, filled it with more than a tonne of soil and then let a pet mouse dig away.

"This thing was bursting at its seams and held together with duct tape," says the evolutionary biologist, now at Harvard University in Cambridge, Massachusetts. "But it worked." It allowed her to study the genetics of burrowing behaviour in a controlled setting. Armed with plastic casts of the burrows and state-of-the-art sequencing, Hoekstra's team discovered clusters of genes that partly explain why the oldfield mouse (*Peromyscus polionotus*) builds elaborate two-tunnel burrows, whereas its close relative, the deer mouse (*Peromyscus maniculatus*), goes for a simple hole in the ground[1].

The findings highlight an underappreciated benefit of a genomics revolution that is moving at breakneck speed. Thanks to cheap and quick DNA sequencing, scientists interested in the genetics of behaviour need not limit themselves to a handful of favourite lab organisms. Instead, they can probe the genetic underpinnings of behaviours observed in the wild, and glean insights into how they evolved. "In my mind, the link between genes and behaviour in natural populations and organisms is the next great frontier in biology," says Hoekstra.

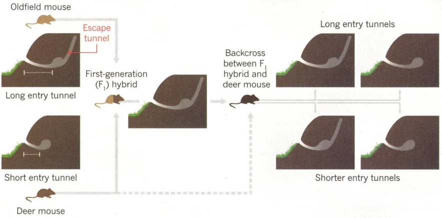
Oldfield mice are native to the southeastern United States, where they burrow in soils ranging from sandy beaches to silt-rich clays. Wherever they dig, their holes look much the same, with a long entrance tunnel and a second tunnel that stops short of the surface and allows them to escape predators. Such invariability hints that the trait is encoded in DNA, says Hoekstra.

To find out where, she and her Harvard colleagues Jesse Weber and Brant Peterson crossbred oldfield mice with deer mice, whose burrows are shallow and lack escape routes. The offspring continued to build complex tunnels, suggesting that the oldfield burrowing genes were dominant (see 'The genetics of burrowing').

A second round of breeding between the first-generation crosses and deer mice revealed that genes linked to burrow length were distinct from those influencing the escape tunnel. Some offspring produced short tunnels with escape routes, whereas others produced long tunnels without them. DNA analysis revealed that three genetic regions are responsible for much of the variation in tunnel length, and a fourth affects escape-tunnel digging.

THE GENETICS OF BURROWING

Mouse cross-breeding shows that distinct gene clusters control the length of entry tunnels and whether or not the mice dig escape tunnels.



PAUL JACKMAN/NATURE

"This paper is awesome," says Cornelia Bargmann, a neurogeneticist at Rockefeller University in New York, noting that it combines cutting-edge molecular-genetics tools with established cross-breeding techniques to study behaviours that have been observed for more than a century in the wild. In the past, geneticists interested in unravelling behaviour had to focus on lab animals for which mutant and transgenic strains and genetic data were available, she says. "But there were always questions we knew would be more interesting in wild animals." Bargmann and her team studied various wild strains of *Caenorhabditis elegans* flatworms to identify genes and brain circuits involved in seeking out new sources of food[2].

> ***"The link between genes and behaviour in natural populations is the next great frontier."***

Cheap genome sequencing is the foundation of this new behavioural genetics. In a paper published on *Nature*'s website this week[3], a team led by Laurent Keller at the University of Lausanne, Switzerland, sequenced hundreds of fire ants (*Solenopsis invicta*) and identified a region that explains why some colonies have only one queen whereas others tolerate more. The region contains hundreds of genes, and Keller thinks that vast behavioural differences among other insects might be explained by such 'social chromosomes'.

These studies offer clues to the evolution of behaviours seen in the wild, says Catherine Peichel, an evolutionary geneticist at the Fred Hutchinson Cancer Research Center in Seattle, Washington. In unpublished work, Peichel and her postdoc Anna Greenwood have discovered a gene that explains differences in schooling behaviour in stickleback fish. The marine variety swims in tight schools, whereas its freshwater cousin swims solo. Such behaviours were probably selected for by different environments — predator-filled seas with few hiding places versus vegetation-filled streams. And Hoekstra says her findings suggest that complex burrowing in mice evolved piecemeal, by combining the genetic modules responsible for simpler digging behaviours.

Hoekstra's team has already found that the key regions for burrowing include genes involved in addictive behaviour. "When you watch these mice, you can't help but think they look like they're addicted to burrowing," Hoekstra says. "You drop one of these oldfield mice in one of our burrowing boxes and they immediately start digging."

The ultimate proof of any gene's role will be to engineer deer mice to build complex burrows by inserting oldfield mice genes, says Hoekstra. Those experiments would have been unthinkable a few years ago, when genome engineering was possible in only a handful of mammals, but this is now conceivable, says Bargmann (see *Nature* **455,** 160–164; 2008). ■

1. Weber, J. N., Peterson, B. K. & Hoekstra, H. E. *Nature* **493,** 402–405 (2013).
2. Bendesky, A., Tsunozaki, M., Rockman, M. V., Kruglyak, L. & Bargmann, C. I. *Nature* **472,** 313–318 (2011).
3. Wang, J. *et al. Nature* http://dx.doi.org/10.1038/nature11832 (2013).

G. M. B. AKASH

More children will be receiving injected killed polio vaccine in a bid to finally eradicate the virus.

HEALTH

Vaccine switch urged for polio endgame

Inactivated virus vaccine could deliver the final blow.

EWEN CALLAWAY IN KADUNA CITY, NIGERIA

By sunrise on a warm December morning, Janila Shulu's team are out in the dirt roads and alleyways of Ungwan Rimi, a poor neighbourhood in a predominantly Muslim section of Kaduna city in northern Nigeria. Three female health workers, accompanied by a community leader, dart from house to house, squeezing a few drops of polio vaccine into the mouths of all the young children they can find, even those who pass by on the street. By 1 p.m., after giving hundreds of doses, they stop for the day — the first of a national five-day effort.

Such campaigns are the backbone of the global push to eradicate polio, but this month the World Health Organization (WHO) in Geneva, Switzerland, proposed a shift in vaccination strategy from oral vaccines to injected ones that may have to be administered in clinics. The change is needed to mop up the last remaining pockets of polio, but experts say that it poses challenges in places such as Kaduna city, which have poor access to health care.

The new policy is an important step towards eradication, says Nicholas Grassly, an epidemiologist at Imperial College London, but implementing it will be difficult. "There are some big ifs as to whether it can happen," he says.

Jonas Salk is credited with developing the first polio vaccine in 1955, an injected vaccine containing killed virus, but the oral live vaccine devised a little later by his competitor Albert Sabin is the workhorse of the Global Polio Eradication Initiative. This public–private effort, started in 1988 and coordinated by the WHO, has cost about US$8 billion so far. The Sabin vaccine is composed of three live but crippled strains of polio. It is cheap and easy to administer, making it ideal for national campaigns that involve tens of thousands of minimally trained workers.

But the live viruses in the Sabin vaccine can revert to disease-causing forms, especially in populations where immunity is not widespread. Northern Nigeria has been battling such vaccine-derived outbreaks since 2005, and one emerged last year in Pakistan (see *Nature* **485,** 563; 2012).

➲ NATURE.COM
For more on the challenges facing vaccine campaigns:
go.nature.com/a3nzqx

In a 4 January announcement, the WHO called for oral polio vaccine containing the polio strain type 2, one of the Sabin vaccine strains, to be phased out — perhaps in as little as two years. The wild form of type 2 has been stamped out globally, but vaccine-derived forms of the strain still circulate in Nigeria and neighbouring countries. Oral polio vaccination will continue, but it will use a vaccine that protects against just the two other types of polio virus that are still circulating in their wild form in Nigeria, Pakistan and Afghanistan.

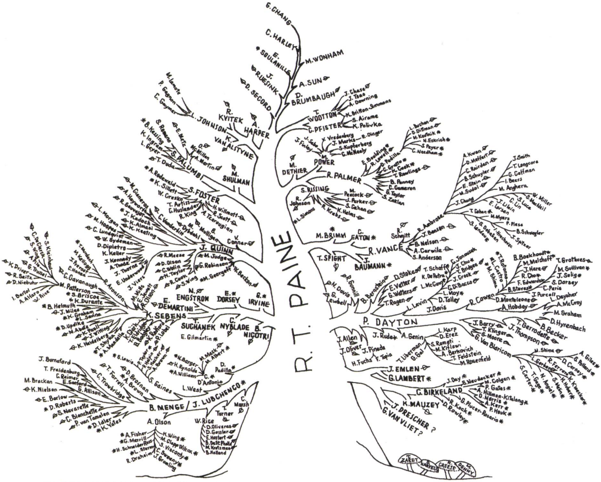
Meanwhile, the policy also calls for the introduction, as quickly as possible, of the oral vaccine's old competitor: the inactivated Salk vaccine. That costs more than ten times as much as the oral vaccine and requires trained health workers to administer it, says Roland Sutter, a vaccinologist at the WHO. But it carries no risk of causing polio. By giving children an inactivated vaccine that protects against all three subtypes of polio, health workers hope to gradually stamp out vaccine-derived outbreaks.

"You have to have a transition period" in which both oral and inactivated vaccines are used, "because if you stop cold turkey you're going to have outbreaks", says Vincent Racaniello, a virologist at Columbia University in New York city. Once the remaining wild polio types are wiped out, the WHO will phase out all oral polio vaccines.

The high cost of the inactivated polio vaccine remains a significant hurdle for the plan, which depends on a reduction in cost to less than 50 cents per dose from the current cost of more than $2, says Sutter. Boosting the immune response by including adjuvants, and delivering the vaccine under the skin instead of into muscle, could help to lower the dose required and cut costs, as could new kinds of vaccine, he says.

Health infrastructure poses another big hurdle, says Grassly. Delivering the vaccine in clinics instead of door to door will pose a challenge for Nigeria, which has one of the lowest rates of routine immunization in the world. Less than 50% of children receive a complete schedule of childhood vaccinations, and in parts of northern Nigeria that figure is around 10%.

"We as a global community have to do a much better job of integrating polio and routine immunization," says Zulfiqar Bhutta, an immunization expert at Aga Khan University in Karachi, Pakistan, and a member of the WHO committee that issued the new vaccination policy. He sees the eventual switch to inactivated vaccines as an opportunity to align polio eradication with routine immunization. "We should have done this a lot earlier," he says. ■


R. T. PAINE
G. CHANG
C. HARLEY
E. SOULANILLE
J. RUESINK
D. SECORD
M. WONHAM
A. SUN
D. BRUMBAUGH
T. WOOTTON
C. PFISTER
M. POWER
M. DETHIER
R. PALMER
S. RISSING
M. BRIMM
C. EATON
R. VANCE
T. SPIGHT
J. BAUMANN
P. DAYTON
J. EMLEN
G. LAMBERT
G. BIRKELAND
K. MAUZEY
J. DRESCHER ?
G. VANVLIET ?
R. KVITEK
A. HARPER
L. JOHNSON
K. VAN ALSTYNE
S. PALUMBI
M. SHULMAN
S. FOSTER
J. QUINN
E. ENGSTRON
E. DORSEY
G. IRVINE
E. DEMARTINI
K. SEBENS
T. SUCHANEK
C. NYBLADE
B. NICOTRI
B. MENGE / J. LUBCHENCO
A. Olson
ZARET
HARVELL
CREESE
DENNY

Main branches with capitalized names indicate Paine graduate students. Secondary branches and twigs with lower case names are second- and third-generation students.
= Ph.D. ✱ = M.S. No sign = current student.
Mushrooms indicate postdocs.

DYNASTY

Bob Paine fathered an idea — and an academic family — that changed ecology.

BY ED YONG

Bob Paine's academic family tree, drawn by Marian Kohn in 1999.

Bob Paine is nearly 2 metres tall and has a powerful grip. The ochre sea star, however, has five sucker-lined arms and can span half a metre. So when Paine tried to prise the creatures off the rocks along the Pacific coast, he found that his brute strength simply wasn't enough. In the end, he resorted to a crowbar. Then, once he had levered the animals up, he hurled them out to sea as hard as he could. "You get pretty good at throwing starfish into deeper water," he says.

It was a ritual that began in 1963, on an 8-metre stretch of shore in Makah Bay, Washington. The bay's rocky intertidal zone normally hosts a thriving community of mussels, barnacles, limpets, anemones and algae. But it changed completely after Paine banished the starfish. The barnacles that the sea star (*Pisaster ochraceus*) usually ate advanced through the predator-free zone, and were later replaced by mussels. These invaders

MARIAN A KOHN

crowded out the algae and limpets, which fled for less competitive pastures. Within a year, the total number of species had halved: a diverse tidal wonderland became a black monoculture of mussels[1].

By re-engineering the coastline in this way[1], Paine dealt a serious blow to the dominant view in ecology of the time: that ecosystems are stable dramas if they have a diverse cast of species. Instead, he showed that individual species such as *Pisaster* are prima donnas, whose absence can warp the entire production into something blander and unrecognizable. He described these crucial creatures, whose influence far exceeds their abundance, as keystone species, after the central stone that prevents an arch from crumbling. Their loss can initiate what Paine would later call trophic cascades — the rise and fall of connected species throughout the food web. The terms stuck, and 'keystone' would go on to be applied to species from sea otters to wolves, grey whales and spotted bass.

Today, ecology students take these concepts for granted — but they shook the field when Paine first articulated them in the 1960s. "He's been one of the most influential ecologists in the last half century," says Simon Levin, a mathematical ecologist at Princeton University in New Jersey, and one of Paine's closest friends. The revelation that not all species are equal was as disruptive to ecology as the loss of *Pisaster* was to Makah Bay. So was Paine's insistence on tinkering with nature — what some have called kick-it-and-see ecology — at a time when most ecologists simply observed it.

But Paine — an organism whose disproportionate influence equals that of any starfish or sea otter — has also changed the ecosystem of scientists. In his five-decade career, he has trained a thriving dynasty of around 40 students and postdocs, many of whom are now leading ecologists themselves and who consider their time with Paine formative. They include Paul Dayton at the Scripps Institution of Oceanography in La Jolla, California, who has shaped understanding of rocky shores, kelp forests and Antarctica's sea floor; Bruce Menge at Oregon State University in Corvallis, who expanded Paine's research to coasts worldwide; Jane Lubchenco, who heads the US National Oceanic and Atmospheric Administration (NOAA) in Washington DC; and Steve Palumbi at Stanford University in California, who used genetics to track the illegal trade of whale and dolphin meat.

"There are other ecologists as famous as Bob, but if you look at their list of students, there aren't nearly as many whom you know by reputation," says Chris Harley at the University of British Columbia in Vancouver, Canada, who was one of Paine's most recent PhD students. Once Paine's students' students are taken into account, his academic family easily stretches into the hundreds. "Everyone is tied to Paine," says Craig McClain, a deep-sea biologist at the National Evolutionary Synthesis Center in Durham, North Carolina, who is three links removed — his postdoctoral adviser was Dayton's student.

> "All my students were smarter than me but just less knowledgeable."

FAMILY VALUES

Science hosts many such dynasties: successions of academic leaders related not by blood, but by mentorship. Each generation inherits attitudes, philosophies and technical skills from the one before. Some, like Paine's, are particularly fertile, sprouting lush branches on the academic tree and driving a field in a new direction. But Paine's dynasty is remarkable not just for its scientific influence, but for its dedicated, tight-knit nature. Thanks to Paine's original — and widely applicable — ideas, his emphasis on independent thought by his protégés and his fun, irreverent nature, almost every member has stayed in science, and specifically in ecology or marine biology.

"It's a surprising list of superstars — great mentors of graduate students, who have published interesting work," says Paine, who retired in 1998 but is still active in the field. These days, Paine can be spotted at ecological meetings by the swarm of academic descendants milling around him. Perhaps in this rich family, there are lessons about why some scientific dynasties flourish and grow, whereas others never bud.

Paine's name is synonymous with coastal life, but his introduction to natural history began on terra firma. As a child in Massachusetts, he went on regular birdwatching walks with a neighbour, who insisted that he record everything he saw. "That was extraordinarily good training," says Paine; it instilled an appreciation for nature and careful observation.

After studies at Harvard University in Cambridge, Massachusetts, and two years' military service, he got his PhD in zoology from the University of Michigan in Ann Arbor, under the late Fred E. Smith. A qualified ecologist in search of an ecosystem, Paine started his own group at the University of Washington in Seattle in 1962 and embarked on his seminal work in Makah Bay. But despite its abundance of starfish, the bay was plagued with marauding beachcombers, who would tamper with Paine's experiments. He craved isolation.

He found it in 1967, after a salmon-fishing trip in the Pacific Ocean, when on a whim he landed on a small island called Tatoosh, just off Washington's Olympic Peninsula. "I just sat around and gawked," he says. "The place was a wonderland of pattern." He saw hundreds of species jostling for space and dominance. He saw starfish prising open mussels, and predatory sea snails drilling into barnacles. He saw whales and sea lions swimming offshore, while raptors and seabirds hovered overhead. And he saw … no one else. With the island protected by the local Makah tribe, there was no one around to ruin his experiments. "I said: this is where I'm going to work."

ISLAND RETREAT

Paine repeated his starfish-clearing experiment on Tatoosh, and saw the same events unfold on a bigger scale: the loss of *Pisaster* triggered a black landslide of mussels that crushed its way down 40 metres of coast. Elsewhere, in a zone dominated by brown algae, Paine systematically added or excluded seven grazing animals by building rings of paint and putty[2]. Two of the species — a sea urchin and a mollusc called a chiton — would annihilate the algae when present in high numbers. The others, all molluscs, had no effect. "It was a colossal effort," says Paine, and the first time that anyone had quantified the influence of so many species in a community. It showed that most are weak interactors, whose absence goes unnoticed. Only a few — including keystone species — are strong interactors that can radically reshape their world. "It was a starting place for untangling the complexity of interactions," says Paine. "If all species were created equal, you wouldn't know where to start."

Paine's views on experimental ecology rapidly percolated through the field, and attracted waves of eager students. "It was a pioneering stage in the field," says Menge. "We felt that we were really the first ones to be doing these sorts of experiments." Tatoosh gave them a place to experiment, and most of Paine's protégés have done tours on its weather-beaten terrain. They camped in garage-sized buildings, abandoned from the island's days as a coastguard outpost. There was no running water and, until a decade ago, no electricity. "It was a brutal environment, and I was out there suffering with them," says Paine. He kept the atmosphere spartan, splashing out only on heavy-duty clothing and boots — the "Paine lab uniform".

Unlike many principal investigators, who dole out predefined projects to new recruits, Paine encouraged freedom and individuality. "He was hands-off until it was necessary to be hands-on," says Menge. Students walked to disparate corners of the island to explore their own passions. One tagged larval rockfish; another studied barnacle dynamics. In the evenings, the group traded data over campfires. "You'd have a day of working the shoreline, you'd trudge back up to the island, and he invariably asked: What did you learn today?" says Tim Wootton from the University of Chicago, Illinois, who studied under Paine in the 1980s. Paine set a high intellectual bar, and was gruff and challenging. "He's very brusque and abrupt, and he kind of scowls a lot," says Lubchenco. But he also encouraged irreverence and mischief — he once signed the name of one his students, Peter Kareiva, on Christmas cards full of lascivious limericks and posted them to the world's top ecologists.

Paine treated his students as peers, supporting their endeavours without directing them. (Smith had treated Paine the same way.) When they published, he kept his name off their papers unless he had had a heavy hands-on role in the research — an ethos that seems unthinkable today. "It hurt him a bit," says Kareiva, now one of Paine's closest friends and chief scientist for the Nature Conservancy, an environmental organization based in Arlington, Virginia. "His presence in the literature would be ten-fold if he hadn't done that." Harley adds: "If Bob Paine were reborn into the system as a junior professor right now, he would turn out excellent graduate students but I don't know how many grants he'd get."

Soon, Paine's students were growing up and embarking on careers of their own. Few have spawned as rich a legacy as Jane Lubchenco and Bruce Menge. They met as graduate students in Paine's lab in 1969, married two years later and began a partnership that has generated more than 31 students and 19 postdocs. After the pair left Paine's lab, they took his experimental approach to the US east coast; she focused on plants and herbivores, while he concentrated on predators. By enclosing, excluding and removing species at different points along the New England shore, they showed[3] that fierce waves can keep predators such as starfish at bay, allowing mussels to dominate. But in sheltered areas, predators kept mussels under control, allowing Irish moss (*Chondrus crispus*), a type of red alga, to take over. The work revealed how the environment can control interactions between species.

L TO R: ANNE PAINE; OREGON STATE UNIV.

Bob Paine with starfish on Tatoosh island, Washington, in 2011 (left); Jane Lubchenco and Bruce Menge with students Eric Sanford and Tess Freidenburg, 1997.

GROWING ECOSYSTEM

After divvying up the New England food web between them, Lubchenco and Menge also split a tenure-track faculty position when they returned west to Oregon State University in 1976. This move, unprecedented at the time, allowed them to further their careers while spending time with their growing family. (One of their children, Duncan Menge, accompanied the Paine clan on field trips and this year is setting up his own ecology lab at Columbia University in New York — a blood child of the academic family.)

The students of the cheerfully described "Lubmengo lab" benefited from the yin-yang qualities of their mentors. Menge was impulsive and had an open-door policy. Lubchenco was deliberate and required appointments. Today, he is laid-back and wears Hawaiian shirts; she is intense and sharp-suited. "The combination of those two was really powerful," says Steve Gaines from the University of California, Santa Barbara, who was one of their star students. "You got an educational experience that would be hard for any individual to give you."

Lubchenco and Menge wanted to probe other environmental impacts on coastal ecosystems, and so began a decades-long world tour that took Paine's experimental approaches to Chile, South Africa and New Zealand. Their studies showed[4], for example, that the intermittent upwelling of nutrient-rich water can intensify competition, predation and other interactions between marine species.

But they also moved — to Paine's disapproval — from his solitary style to large teams. Their grandest venture is the Partnership for Interdisciplinary Studies of Coastal Oceans (PISCO) — a 13-lab collaboration that began in 1999. The participants run regimented studies along a 1,900-kilometre stretch of coastline from Alaska to Mexico, including a census of local marine life and measurements of water conditions. There are some Paine-style manipulations, but PISCO is more about collecting data to reveal how the oceans are changing. When Oregon's beaches started filling with dead crabs and fish in 2002, PISCO's data quickly revealed the cause — an oxygen-deprived dead zone that had been mysteriously growing in the water off the coast (see *Nature* **466,** 812–814; 2010). "The problems we're trying to solve in ecology are way beyond the local scale," says Menge. "My dream is for PISCO-like entities in all the marine ecosystems around the world."

But that is not Paine's dream: big science is an anathema to him. "He was always grousing about how this massively expensive work was really not much more than stamp-collecting," says former student Richard Palmer, now at the University of Alberta in Edmonton. Paine fears that PISCO's approach will yield broad, unclear trends rather than detailed insights. Worse, it risks robbing students of the playful creativity that he tried to instil. "They're so involved with making the same measurements up and down the same kilometres of coast," he laments. "My loosey-goosey attitude to Tatoosh was to get brilliant people to do what they want to do." But Lubchenco and Menge argue that ecology's challenges, from global warming to ocean acidification, are outpacing the rhythms of solo experiments. "Bob fought that a lot," says Lubchenco. "People had to rebel against their parent."

Lubchenco's move into policy was similarly rebellious. She served as president of the American Association for the Advancement of Science in 1997 and of the Ecological Society of America between 1992 and 1993. And as administrator of NOAA from 2009, she has influenced the US government's response to the Deepwater Horizon oil spill, and its strategy to manage the oceans sustainably. "The culture we grew up with was very anti being relevant," says Lubchenco, who will leave NOAA at the end of

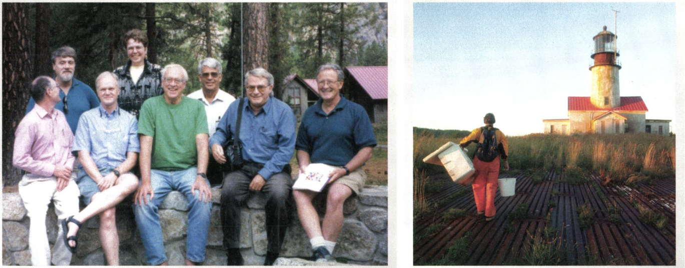
➲ NATURE.COM
For more about Bob Paine's dynasty, listen to the podcast:
go.nature.com/olfjxo

February to return to research. "I knew Bob didn't approve and I did it anyhow. It was really painful to rebel."

Any such disapproval has long since vanished, and Paine now speaks of Lubchenco's influence with deep respect. He has reluctantly become a key player in conservation himself, leading panels of scientists who have assessed the decline of the Steller's sea lion (*Eumetopias jubatus*) and the recovery of Alaskan waters following the 1989 *Exxon Valdez* oil spill. "My role was simply to oversee the mob," he says. "I was bullied into it."

NEXT GENERATION

Lubchenco's academic children — Paine's grandchildren — have been steeped in 'relevance' from the start. "We were all surrounded by this environment where it was encouraged to think about steps connecting the science to policy," says Heather Leslie at Brown University in Providence, Rhode Island, a former Lubchenco student.



L TO R: COURTESY OF BRUCE MENGE & JANE LUBCHENCO; MATTHEW RYAN WILLIAMS/*THE NEW YORK TIMES*/REDUX/EYEVINE

Paine (front row, centre) surrounded by colleagues and protégés at his retirement party in 1999 (left); Paine's former student Cathy Pfister on Tatoosh in 2012.

Gaines epitomizes this approach. The ultimate blend of the Paine and Lubmengo heritage, he is an experimental ecologist who uses basic science to influence policy, is not shy of collaborations and has won a national award for mentoring. "Steve sits right at that sweet spot in many different ways," says Lubchenco. Whereas Paine studies how species interact on individual shores, Gaines looks at connections across entire oceans. He has shown[5] that the larvae of fish and other marine life sometimes drift for hundreds of kilometres on ocean currents, even if the adults are fixed in place. "When Bob removed the sea star and got an explosion of mussels, those mussels were coming from somewhere else," he says. These planktonic drifters connect disparate parts of the ocean, with huge implications for marine reserves, where fishing is forbidden. Rather than reducing fish catches, Gaines showed[6] in 2005 that these reserves have the potential to make nearby fisheries more productive as their larvae disperse and replenish the stocks.

Gaines' own students, in the style of the entire family, have branched out into new areas. Kate Smith, who did her PhD with him and is now at Brown, applied Gaines' ideas on dispersing larvae to understand how infectious diseases disperse. Her ecological interests are the same as those of her academic parents and grandparents — where organisms are, why they are there and how they influence each other — but applied to microbes and continents, rather than invertebrates and tidal pools. "It's all related to connectivity," says Gaines.

There are other ways than Paine's to spawn a dynasty. Take, for example, the chain of biological luminaries that emerged in the 1940s, beginning with physiologist James Shannon at the US National Institutes of Health. His descendants — including Steve Brodie, who did pioneering work in drug metabolism and Julius Axelrod, who was awarded a Nobel prize for his work on the release of neurotransmitters — helped to revolutionize human pharmacology and neuroscience. "Their approach was: Don't feel you have to have all your i's dotted and t's crossed. Just go and do it," says Robert Kanigel, a writer who chronicled the dynasty in his book *Apprentice to Genius* (Macmillan, 1986).

But the Shannon lineage was radically different from Paine's. There was no premium on independence; instead, students served as apprentices to their masters. "In many of the key experiments, the younger person was doing the work of the mentor, and that sometimes led to resentments," says Kanigel. Such relationships can breed envy and mistrust, he says. "People wonder if they'll get the credit they deserve."

Paine proves that the opposite strategy works. "Treat your graduate students as human beings and be accessible to them. That sense of social equality is very important," he says. "All my students were smarter than me but just less knowledgeable." This attitude selected for self-propelled, passionate students who could find their own way. "It helped in getting the experience to implement your own research programme," says Wootton.

Now aged 79, Paine is still conducting research and inspiring students. He joined a crop of them last summer for a week in Patagonia. "He'd be in the intertidal on his hands and knees and asking questions," says Kareiva, who was also on the trip. Paine also makes regular excursions to Tatoosh, where research is now run by Wootton and his wife, Cathy Pfister, who is also at the University of Chicago. Paine's mind is as sharp as ever, but having lost good binocular vision, his steps are less sure. "I hire my very athletic daughter to haul me around the rocks," he says. "It's not sufficient, but I can continue to keep track of my long-term experiments."

In 1995, Paine finally allowed starfish to return to the Tatoosh shore after 25 years of continuous exile. "I thought, I'm not going to live forever, so let's see what happens if I let the starfish back." Their numbers have rebounded, the mussels are losing ground, and the shore is returning to the state it was in before Paine's interference. The same cannot be said for ecology. It will never be the same after Paine. ■

Ed Yong *is a freelance writer and author of the blog Not Exactly Rocket Science, based in London.*

1. Paine, R. T. *Am. Nat.* **100,** 65–75 (1966).
2. Paine, R. T. *Nature* **355,** 73–75 (1992).
3. Lubchenco, J. & Menge, B. A. *Ecol. Monogr.* **48,** 67–94 (1978).
4. Menge, B. A. *et al. Ecol. Res.* **17,** 1–16 (2002).
5. Kinlan, B. P. & Gaines, S. D. *Ecology* **84,** 2007–2020 (2003).
6. Gaylord, B., Gaines, S. D., Siegel, D. A. & Carr, M. H. *Ecol. Appl.* **15,** 2180–2191 (2005).

TORU YAMANAKA/AFP/GETTY

FALLOUT OF FEAR

AFTER THE FUKUSHIMA NUCLEAR DISASTER, JAPAN KEPT PEOPLE SAFE FROM THE PHYSICAL EFFECTS OF RADIATION — BUT NOT FROM THE PSYCHOLOGICAL IMPACTS.

BY GEOFF BRUMFIEL

The first thing Kenichi Togawa does when he comes home from work is switch on his video-game console. The 39-year-old father of three spends hours each evening playing video games and drinking *shochu*, a strong Japanese liquor. He often falls asleep in front of the television, then wakes up shivering and crawls into bed with his wife, Yuka.

For nearly two years, Kenichi and his family have been refugees from the worst nuclear disaster in 25 years. On 11 March 2011, a giant earthquake struck off the northeast coast of Japan, sending a 13-metre-high wall of water into the Fukushima Daiichi nuclear power station and triggering meltdowns in three of the six reactors. The next day, just hours before the Unit 1 reactor exploded, the Togawa family fled their home 10 kilometres from the plant. Today, they live in a tiny flat outside the evacuation zone — one of dozens in a series of slate-grey temporary buildings in the northeast section of Fukushima prefecture. The five Togawas are bundled into three rooms totalling just 30 square metres, with windows poorly insulated against the winter winds.

The past 18 months have taken a mental toll on the family. Kenichi, who had worked at the nuclear plant, was once a keen judo fighter who went out often with friends, but the radioactivity has scattered his martial-arts club. These days, he exercises less and rarely socializes. He drinks more and has put on weight.

Yuka is prone to public outbursts of anger, unusual among Japanese women in the relatively traditional Fukushima prefecture. She is happy when she thinks about day-to-day life, but when her mind turns to the long term, as it inevitably does, she feels depressed. "This is temporary," she says. "We leave our house in the morning and we come home and it's temporary. It's like floating in the air."

Other people they know are struggling even more. Many of their current neighbours are out of work and stay at home all day. Some of Kenichi's former colleagues sent their wives and

A woman prepares lunch in her little partitioned unit at an evacuation centre, Miyako April 2011.

children away, from fear of radioactive contamination, while they stayed to work.

In the immediate aftermath of the nuclear accident, public-health experts worried about the possible risk from radiation. Subsequent analyses have shown that the prompt, if frantic, evacuation of areas around the reactors probably limited the public's exposure to a relatively safe level (see 'The evacuation zones'). But uncertainty, isolation and fears about radioactivity's invisible threat are jeopardizing the mental health of the 210,000 residents who fled from the nuclear disaster.

Researchers and clinicians are trying to assess and mitigate the problems, but it is unclear whether the Japanese government has the will, or the money, to provide the necessary support. Nor is it certain that the evacuees will accept any help, given their distrust of the government and their reluctance to discuss mental problems. This combination, researchers fear, could drive up rates of anxiety, substance abuse and depression.

The nuclear evacuees face a more difficult future than the survivors of the tsunami, which left nearly 20,000 dead or missing and caused billions of dollars in damage. "The tsunami-area people seem to be improving; they have more positive attitudes about the future," says Hirooki Yabe, a neuropsychiatrist at Fukushima Medical University, who has been working with both groups. Nuclear evacuees "are becoming more depressed day by day".

ESCAPE

Fukushima prefecture is a patchwork of orchards, rice paddies and fishing villages. In the 1970s and 80s, coastal residents welcomed nuclear power and two plants were built to supply electricity to Tokyo. Kenichi started working at Fukushima Daiichi in 1994, and at the time of the accident was a contract maintenance engineer. Yuka worked as a hospital nurse. The Togawas and their children, now aged 9, 12 and 15, lived in a four-room flat in Namie, a small, close-knit seaside town.

The family's life was upended at 14:46 on 11 March 2011. Kenichi was in the smoking room at the plant when he felt the ground shudder for several minutes. He ran back to his office, weaving around scattered desks and downed ceiling panels, and grabbed his driving licence and car keys. But he quickly found that traffic had clogged the routes out of the plant because the quake and tsunami had destroyed bridges and roads. Kenichi ditched his car and walked the remaining 8 kilometres home.

There he found that all his family members were safe, but he worried about the plant. In his job, he had overseen maintenance of the systems meant to cool the reactor in an emergency. If those had failed, he knew that a meltdown would soon follow, spreading radiation to nearby towns. That night, as aftershocks rocked the house, his family slept fitfully with the lights and television on.

Kenichi was right to be alarmed. The tsunami had knocked out the generators that pumped cooling water into the reactor cores. As temperatures rose, the slender fuel rods full of uranium pellets began to warp. The meltdowns had begun.

Early the next morning, a siren wailed across Namie, signalling an evacuation. The Togawas were told to move to Tsushima, 30 kilometres to the northwest. After retrieving their car, the family set out, but the roads were choked with panicked residents and the Togawas ended up at a different evacuation centre. When Kenichi learned that the emergency diesel generators at the plant had also failed, he bundled the family into the car once more, hoping to reach Tsushima. "We have to run away," he remembers thinking in a panic.

On the way, Kenichi received a text message from a friend who worked at the power company's offices in Tokyo. Unit 1 had exploded, and radioactivity was spreading across Fukushima. The family drove from one full evacuation centre to the next, until they reached a dark, cramped gymnasium in Kawamata, around 40 kilometres northwest of the plant. There they were given a small patch of hardwood floor to call home. But they were still deeply worried about the radiation. "We didn't know much about radiation's effects, and we didn't know if Kawamata was safe or not," says Yuka.

Japan is used to natural disasters, and immediately after the tsunami hit, the country's emergency services sprang into action. Groups of doctors and emergency workers from around the nation struck out for the northeast coast to begin search and rescue operations and to administer care. The medical university in Fukushima City became a hub. In the days and weeks after the accident, the university hospital took in seriously ill patients from the coast. It also found itself on the front line of the nuclear emergency: doctors used Geiger counters to screen evacuees' thyroid glands, which are particularly sensitive to radiation, and treated several workers from the nuclear plant, who had suffered high radiation exposures.

FIRST RESPONDERS

Mental-health experts were among the first responders, reflecting an ongoing change in Japan's attitudes towards mental health. For many years, Japan's modest but modern mental health services were geared to help only the most severely mentally ill. The society has traditionally paid little attention to more routine disorders such as depression. In recent years, however, the Japan Medical Association has started educating doctors about depression and suicide, and the national government has conducted public suicide-prevention campaigns.

Still, the quality of care remains patchy, and even before the accident, Fukushima prefecture was not a bright spot. Mental health was a not a priority for the rural, conservative region or its taciturn citizens. As a result, the tsunami and nuclear disaster strained the region's mental-health services to near breaking, says Yabe.

In the wake of the accident, most of the prefecture's resources were devoted to helping

EMPTY STREETS: In the aftermath of the nuclear disaster, 210,000 people had to evacuate Tomioka and other towns in Fukushima prefecture. Nearly 156,000 remain displaced from their homes.

OSAKABE YASUO/DEMOTIX/CORBIS

those with established mental disorders. Yabe, for example, packed his car with antipsychotic and anticonvulsive medication and made runs to Soma City, where many evacuees had ended up. Mental-health professionals visited the cramped shelters elsewhere, but they tended to treat only the most severe cases of delirium and post-traumatic stress disorder.

The Togawas were among thousands of people left to their own devices by the overwhelmed doctors and counsellors. The family's first days in the cramped shelter are difficult to recall now, says Yuka, but what she can remember isn't pleasant: ill and elderly patients lying on the floor; ongoing fear about radiation; evacuees jumping queues and snatching food. "We were like dogs and cats without chains," she says.

With little guidance from the outside, the shelter's residents tried to organize themselves. Yuka volunteered her nursing skills, but after working for three days, she was filled with anger: why should she, a victim, have to spend all her time helping others, she wondered. Yuka locked herself in the family's car outside the shelter, "and just exploded and screamed and shouted and cried".

SUBTLE DAMAGE

As the evacuees struggled to adjust, so too did the doctors and psychologists at Fukushima Medical University. By May, the emergency response was mostly over and the hospital had a new job — to assess the public's radiation dose. The task has proved tricky, says Shunichi Yamashita, a radiation health expert at Nagasaki University, who was brought in to head the Fukushima Health Management Survey. The radiation monitors around Daiichi were damaged or destroyed by the earthquake and tsunami, and the chaotic nature of the evacuation makes it difficult to assess how long and severely each person was exposed.

The few attempts made so far, however, have generally shown minimal risk. The health survey's latest assessment suggests that the dose for nearly all the evacuees was very low, with a maximum of only 25 millisieverts (mSv), well below the 100-mSv exposure that has been linked to an increased risk of cancer in survivors of the atomic bombs dropped on Hiroshima and Nagasaki in 1945. The World Health Organization also issued a reassuring report in May, saying that most evacuees from places like Namie received estimated doses between 10 and 50 mSv. It did note, however, that infants might have received a dose that could increase the risk of cancer in their still-developing thyroids.

Radiation specialists say that it is difficult to predict the health effects from such low doses. "I think it's likely that there will be increased cancer risks, but they will be very, very small," says Dale Preston, an independent statistician who has studied atomic-bomb survivors. "If you did a large study, I think your chance of observing a statistically significant radiation-associated risk would be pretty low."

With that in mind, the health survey decided against following a fixed cohort to study the incidence of disease. Instead, it provides thyroid screening and other health checks to any evacuees who desire them. The hope is that the screenings themselves, along with the data collected, will help to reassure the public that the risks are low, says Yamashita.

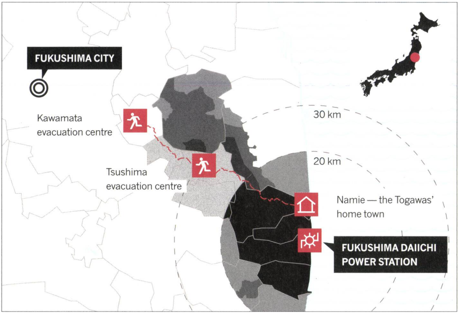
Mental health has been a major component of the survey. In January 2012, researchers sent out questionnaires to all 210,000 evacuees to assess their stress and anxiety. The levels tabulated among the more than 91,000 respondents were "quite high", says Yuriko Suzuki, a psychiatrist at the National Institute of Mental Health in Tokyo. Roughly 15% of adults showed signs of extreme stress, five times the normal rate, and one in five showed signs of mental trauma — a rate similar to that in first responders to the attacks of 11 September 2001 in the United States. A survey of children, filled out by their parents, showed stress levels about double the Japanese average.

The stress has pushed some evacuees to breaking point. On a crisp day last November, Kenji Ookubo wandered through Iitate, a village 40 kilometres northwest of the plant, practising his golf swings in the empty streets. The town had been evacuated after the accident because it lay in the path of the plume of radiation blowing away from the plant. But Ookubo couldn't stand the temporary housing, where he had started drinking and suffered from stomach aches. After renting a room in Kawamata, he began squatting in his parents' abandoned home. "I came back just to run away from the stress," he says. With no job, and no prospects, "I can't see the future," he says.

It is a pattern seen frequently after major catastrophes, says Ronald Kessler, a professor of health-care policy at Harvard Medical School in Boston, Massachusetts. "In the short term, people get energized," he says. But when extensive damage or health problems prevent them from getting back to their old lives, depression and anxiety set in. "When something this big happens, it's just ridiculously daunting," he says. "At a certain point you just get worn down."

His own surveys of people evacuated after Hurricane Katrina, which struck the United States in 2005, show[1] that property loss and health concerns were the main causes of anxiety. Whereas many survivors of the Japanese tsunami have seen their homes rebuilt and lives restored, nuclear refugees are still dealing with both of those problems. Above all, the fear of radioactivity takes a unique toll. "It's something you don't feel; you don't notice what happened, and yet you understand that there are these long-term risks," says Preston. "It's scary."

Little is known about the long-term effects of


FUKUSHIMA CITY
Kawamata evacuation centre
Tsushima evacuation centre
30 km
20 km
Namie — the Togawas' home town
FUKUSHIMA DAIICHI POWER STATION

THE EVACUATION ZONES: During the first stages of the nuclear crisis at the Fukushima Daiichi power station, Japan evacuated a 20-kilometre radius around the plant, forcing the Togawa family to flee. They stopped at several centres before reaching Kawamata, where they currently reside. Japan later updated and expanded the evacuation zones.

■ Fully evacuated zone.
■ Dose rate of 20–50 millisieverts per year. Entry is restricted.
▩ Dose rate >20 millisieverts per year. Some access allowed. No protective equipment required.
▨ Residents may move freely but not stay over night.
░ Planned evacuation area.

MAI NISHIYAMA

THE TOGAWAS: After nearly two years in temporary housing, the family worries about the future but is happy to be together. From left, Shoichiro, Kenichi, Rina, Yuka (back) and Kae (front).

that fear, in part because nuclear accidents are so rare. But the 1986 disaster at the Chernobyl nuclear power plant in Ukraine suggests that fear of radiation can cause lasting psychological harm. Two decades after the accident, those who had evacuated as children complained of physical ailments more often than their peers, even though there was no difference in health[2]. And the mothers of those children suffered from post-traumatic stress disorder at about twice the rate of the general population, says Evelyn Bromet, a psychiatrist at the State University of New York in Stony Brook. Other studies of Chernobyl's aftermath found that evacuees had elevated rates of depression[3] and that a subset of clean-up workers committed suicide at a rate about 1.5 times that of the general population[4].

For Fukushima evacuees, says Bromet, "There's going to be a tremendous amount of health-related anxiety and it's not going to go away easily."

FEAR FACTOR

Yabe says that "radiophobia" remains a major problem among the Japanese refugees. A poll published last year by the Pew Research Center in Washington DC, for example, found that 76% of Japanese people believed that food from Fukushima was not safe, despite government and scientific assurances to the contrary. And many do not trust the government health surveys that found very few cases of significant radiation exposure among evacuees.

Yuka shares some of those concerns. She and Kenichi have educated themselves, and they have gained some reassurance from regular health checks and thyroid screenings. The children carry dosimeters provided by the health survey to collect radiation data and to calm public concerns. But Yuka wonders whether they will one day develop cancer.

At the moment, however, the family is preoccupied with practical concerns. The government has said that the Togawas can remain in their small flat until August 2014, but after that, Yuka says, they don't know what will happen. "The government officials say that they are working on it and that they are trying to construct public housing for those people who had to evacuate. But where? Nothing is clear." Whenever she and Kenichi think about the long term, they start to feel depressed.

The scientists involved with the Fukushima Health Management Survey have assigned a team of psychiatrists and nurses to make follow-up phone calls to individuals who had high scores for distress on the mental-health questionnaire. But only about 40% of adults responded to the questionnaires, and the researchers suspect that the most severely affected people did not participate. Even when the psychiatrists can connect, the evacuees usually don't stay on the phone for more than five to ten minutes. "Northern people are a very closed people, they don't really talk about their personal things, especially to somebody they've never met before," says Yabe.

Even when the psychiatrists identify problems, it is unclear what to do about them. Most evacuees, like the Togawas, are suffering from sub-clinical problems — mental anxiety and stress that affects their everyday life but does not require hospitalization or extensive therapy. There is no established treatment regime for such survivors from large disasters, says Suzuki.

Yabe suggests that walk-in clinics specializing in mental health could be set up throughout Fukushima prefecture to engage communities and help families. Suzuki says that involving large segments of the population in group-therapy sessions might be the way forward. Many say that it would help for evacuees to develop a sense of community — but the government has not fostered that. Temporary houses are "strung out like a railroad", says Bromet. The government could have built them "in a circle with a playground in the middle, or some obvious place for people to meet, but they didn't", she says.

Kessler says that unlike the tsunami survivors, whose grief will lessen over time, the nuclear evacuees could experience growing anxiety, particularly about radiation. "When everything has settled down, that will be a huge, rife issue," he predicts. Now is the best time to try to get ahead of these problems, he says. "There's a window of opportunity."

➲ **NATURE.COM**
For an interactive graphic and to hear from evacuees, see: go.nature.com/lte49b

But the health survey lacks the funding for a more ambitious programme. The national government has given it just ¥3 billion (US$34 million) a year, but it is currently consuming about twice that amount, so the survey is under enormous financial pressure, says Seiji Yasumura, one of its leaders and an epidemiologist at Fukushima Medical University. So far, only 100 of the 210,000 evacuees have been interviewed face-to-face by mental-health experts.

Little by little, things are getting better for the Togawas. The children seem happy in their new school, and in September 2011 Kenichi found a job with the local government, clearing contaminated soil from the homes of neighbours. "He's worked so much overtime that his company is saying he has to have a break," boasts Yuka. She has found part-time work as a nurse in a local clinic. Her occasional outbursts sometimes cause tension with co-workers, but she enjoys speaking her mind: "I say what I want to say."

After filling out one of the health-survey questionnaires last year, Yuka got a flyer in the post inviting her to talk to someone over the phone. She thought about it but decided not to. "I don't feel like phoning. It's been nearly two years," she says. "I don't know what to say." ■ **SEE EDITORIAL P. 271**

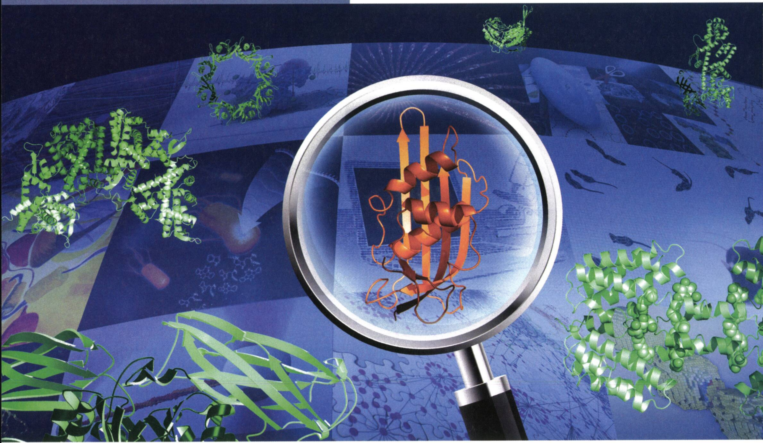
Geoff Brumfiel *is a senior reporter at* Nature.

1. Galea, S. *et al. Arch. Gen. Psychiatr.* **64,** 1427–1434 (2007).
2. Bromet, E. J. *et al. BMC Publ. Health* **9,** 417 (2009).
3. Havenaar, J. M. *et al. Am. J. Psychiatr.* **154,** 1605–1607 (1997).
4. Rahu, K., Rahu, M., Tekkel, M. & Bromet, E. *Ann. Epidemiol.* **16,** 917–919 (2006).

nature | methods

Method of the Year:

Targeted Proteomics



Nature Methods presents a series of articles that report on the technology development and basic and translational applications of targeted proteome analysis. The Methods to Watch feature provides a glimpse at possible future Methods of the Year.

Access the Method of the Year and Methods to Watch by visiting:
www.nature.com/nmeth/focus/moy2012

nature publishing group npg

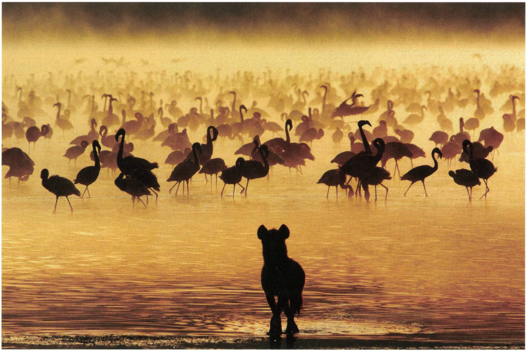
COMMENT

COMMUNICATION Sally Rockey reflects on two years of blogging at the NIH **p.298**

ECOLOGY Zoological travelogue tracks rare species worldwide **p.300**

WOMEN Calls to root out sexism in journals, conferences and experiments **p.305**

OBITUARY Rita Levi-Montalcini, nerve growth factor pioneer and science advocate **p.306**



YI LU/VIEWSTOCK/CORBIS

A hyena surveys a flock of flamingos in South Africa.

Time to model all life on Earth

To help transform our understanding of the biosphere, ecologists — like climate scientists — should simulate whole ecosystems, argue **Drew Purves** and colleagues.

No report from the Intergovernmental Panel on Climate Change would fail to mention global climate models. Yet the international bodies that are charged with addressing global challenges in conservation — including the Intergovernmental Platform on Biodiversity and Ecosystem Services, which holds its first plenary meeting next week in Bonn, Germany — cannot refer to analogous models of the world's ecosystems. Why? Because ecologists have not yet built them.

General circulation models, which simulate the physics and chemistry of Earth's land, ocean and atmosphere, embody scientists' best understanding of how the climate system works and are crucial to making predictions and shaping policies. We think that analogous general ecosystem models (GEMs) could radically improve understanding of the biosphere and inform policy decisions about biodiversity and conservation. Currently, decisions in conservation are based on disparate correlational studies, such as those showing that the diversity of bird species tends to decline in deforested landscapes. GEMs could provide a way to base conservation policy on an understanding of how ecosystems actually work.

Such models could capture the broad-scale structure and function of any ecosystem in the world by simulating processes — including feeding, reproduction and death — that drive the distribution and abundance of organisms within that ecosystem. Ecologists could apply a GEM to African savannas, for instance, to model the total biomass of all the plants, the grazers that feed on the plants, the carnivores that feed on the grazers and so on. Over time, the flows of energy and nutrients could be mapped between them. All of the organisms would be grouped not by species, but according to a few key traits such as ▶

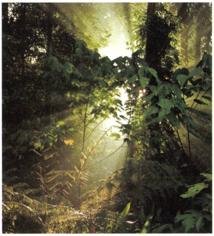
▶ whether they are plants, birds or mammals, cold blooded or warm blooded, diurnal or nocturnal. By encoding processes such as migration and predation into simple mathematical and computational forms, ecologists could model what happens to the various groups over time.

Metrics such as the diversity of animal types inhabiting the grasslands could be used to assess the savannas' health, stability and resilience, and to analyse the fate of particular groups of organisms such as top predators. Ecologists could explore how these attributes might change in response to, say, climate change, the introduction of invasive species or poaching. And, because the rules of play are likely to be broadly similar no matter what the ecosystem, the GEM could equally be applied to forests, lakes or the remotest parts of the ocean, providing a common framework for understanding and managing disparate ecosystems on local and global scales.

There are huge challenges to building GEMs — not least, obtaining the appropriate types of data to validate the models' predictions. But the difficulties are not insurmountable. Theories abound for describing the processes that drive ecosystems, many of which are backed up by data.

BUILDING A PROTOTYPE

Over the past two years, we at Microsoft Research and at the United Nations Environment Programme World Conservation Monitoring Centre, both in Cambridge, UK, have built a prototype GEM for terrestrial and marine ecosystems. Called the Madingley model, it uses real data on carbon flows as a starting point. We have hit all sorts of computational and technical hurdles, and are expecting more as we develop the model. Yet the project demonstrates that building GEMs is possible. From the relationship between the mass of individual organisms and how long they live, or the effects of human perturbations such as hunting, to the distribution of biomass across Earth (see 'Model life'), the model's outputs are broadly consistent with current understanding of ecosystems.



Modelling every organism in an ecosystem such as a tropical rainforest would be impossible.

SIMON LONG/FLICKR/GETTY

Modelling ecological phenomena at various scales is not new. Conservationists frequently use models to predict how much habitat fragmentation an endangered species can tolerate. But few ecologists have tried to build models of how general ecosystem properties emerge from the interactions of individuals. One attempt — known as Ecopath with Ecosim[1] — is being developed at the University of British Columbia's Fisheries Centre to address management and ecological questions. This combines modelling at the phytoplankton scale with that at the level of fisheries and marine mammals. Another, called Atlantis[2], developed by Australia's Commonwealth Scientific and Industrial Research Organisation, incorporates biophysical, economic and social factors to provide an integrated tool for modelling marine ecosystems. But these consider fewer processes and organisms than would a GEM.

Throughout the history of ecology, most researchers have resisted abstraction because ecological complexity is so obvious in nature. Ignoring the myriad shapes and colours of different bird species, for example, seems instinctively wrong. Instead, ecologists have tended to stress the importance of species identification as well as a vast number of ecological processes, from individual adaptation to the social dynamics of groups. Many in the field also emphasize that findings in one ecosystem do not generalize to others, and that randomness and history could be as important in affecting some particular measurement as any deterministic rules[3]. But comprehensive species-specific data will always be in short supply (at least 80% of the millions of species on Earth are undescribed[4]), so a better understanding of ecosystems demands a broad-brush approach.

Building a GEM will require different types of data — to help define the ecological rules at play, to provide reasonable starting conditions for the simulation (such as a realistic ratio of herbivores to top predators) and to evaluate the model's predictions. For these three goals, enough data are available to get started, although information on ecological processes far outweighs the rest.

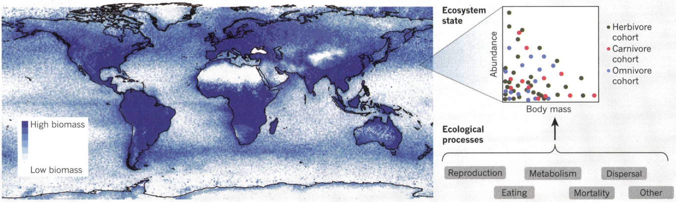
Metabolic rates, for instance, have been measured in hundreds of animals in the lab, and researchers in the field have documented lifespans, growth rates and reproductive success for thousands (in some cases, millions) of birds, mammals, plants and bacteria. Ecologists have also mathematically determined numerous 'rules of existence' for some organisms, such as that an animal's metabolic rate is proportional to its mass raised to a power of around 0.70 (ref. 5).

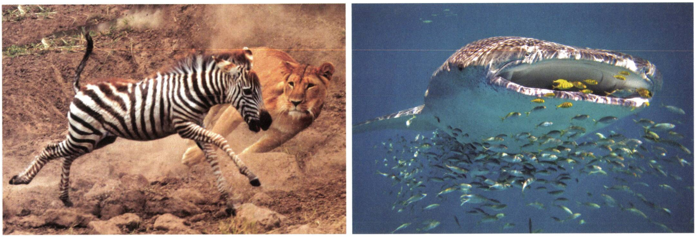
MODELLING BEHAVIOUR

Obviously, modelling every organism within an ecosystem is impossible. (We estimate that it would take a standard laptop computer around 47 billion years to model for 100 years every multicellular animal within just one of the 1-degree grid cells covering

MODEL LIFE

Variation in biomass across the world simulated by the Madingley model for terrestrial and marine ecosystems. Fundamental ecological processes, encoded into simple computational forms, determine the abundance and body mass of organisms (grouped into cohorts for simplicity) and so indicate the state of ecosystems.





ZEBRA: MANOJ SHAH/GETTY; WHALE SHARK: ANDY ROUSE/NATUREPL.COM

From hunting zebra to filter feeding, the process of predation in all ecosystems plays by similar rules.

Earth.) Yet certain computational techniques have been developed, mainly in marine ecology, that could allow researchers to model entire ecosystems using rules about the behaviour of individuals.

One approach is to model collections of organisms or 'cohorts'. The idea is that within a cohort, individuals are similar enough to be considered identical. For a shoal of small herbivorous fish meeting a cloud of plankton, say, ecologists could calculate the feeding rate for an exemplar fish and then apply that rate to the whole shoal. (In the simulation just described, we found that grouping organisms into cohorts on the basis of body size, functional group such as omnivore or carnivore and a few other traits reduced the computation time to 10 hours).

The biggest stumbling block to constructing GEMs (after convincing ecologists that they can and should be built!) is obtaining the data to parameterize and validate them.

Records of what species of plants and animals live in the world's forests, grasslands and oceans are often available to some extent, but far fewer data exist on the abundance of those species. And almost no data have been collected on the properties of whole ecosystems, such as on the distribution of body sizes from plankton to whales. Marine trawl surveys carried out for research or to assess fish stocks probably come closest to providing this type of information, although even these are restricted as to what size range of organisms they survey.

A new programme of data gathering is easy to envisage. Using automated cameras and image recognition, it should be possible to sample thousands of animals and determine their approximate size and what broad group they belong to: reptile or mammal, flying or non-flying. Motion-activated cameras used by conservationists and wildlife enthusiasts already produce tens of thousands of images of fish, birds and mammals every day. And stored away in numerous research institutions are vast samples of insects collected in traps that suck them out of the air, and data from continuous plankton recorders towed beneath ships for millions of kilometres.

Naturally, a major new data-gathering programme would be costly. But globally, governments already spend billions of dollars on satellite observations of vegetation and habitat distribution, fisheries surveys, forest inventories and species surveys. Diverting a small fraction of these funds to gathering the data needed to develop and evaluate GEMs could pay dividends. A first step would be for governments around the world to support programmes similar to the National Ecological Observatory Network — an international cooperation funded by the US National Science Foundation to manage large-scale collection of ecological and climate data.

> *"Ecological systems do not have the equivalent of the precise laws used by climate scientists."*

To reduce costs and to harness the power of citizen science, data collection could even be crowd-sourced. Rapidly growing websites such as iNaturalist and eBird (on which users can share their observations of wildlife) currently focus on traditional species identification. Such sites could potentially collate an extraordinary amount of information on functional groups of organisms and traits such as body size.

TRUSTED ADVICE

Constructing realistic GEMs is one thing. The real challenge is to produce models from which the predictions are trustworthy enough to guide the decisions of conservationists and policy-makers. Recent progress in computational statistical methods offers a way for ecologists to formally build trustworthiness into models. For instance, tools are available to quantify the uncertainty associated with models' predictions. A healthy crop of alternative, competing GEMs will be crucial, together with mechanisms that enable their fair assessment. In blind-testing, for example, different models could be used to predict an ecosystem property that has been measured but not reported, allowing the models to be ranked in terms of how well they do.

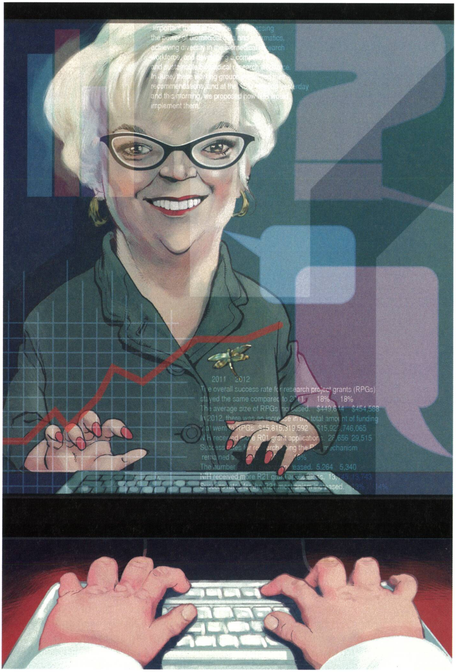
We are not proposing that GEM predictions (which will always be simplistic) provide the only guide to conservation policy and the management of ecosystems. But coupled with models from other fields, such as economics and epidemiology, they could offer a means of managing human actions and the biosphere in an integrated, consistent and evidence-based way. Far from eclipsing traditional ecological research, GEMs would draw on it and give such work more focus. Using GEMs, ecologists could identify processes that are poorly understood yet crucial to ecosystem structure and function, rather than delve deeper into well-studied areas.

Ecological systems do not have the equivalent of the precise laws used by climate scientists. This is a significant challenge to building GEMs, along with the complexity of nature, the small number of GEM-like models under development and the paucity of data with which to constrain them. But just by attempting to build general models, ecologists will find out what they need to know to truly understand ecosystems. ■

Drew Purves *is head of the Computational Ecology and Environmental Science Group at Microsoft Research in Cambridge, UK.*
Jorn Scharlemann, Mike Harfoot, Tim Newbold, Derek P. Tittensor, Jon Hutton, Stephen Emmott.
e-mail: dpurves@microsoft.com

1. Christensen, V. & Walters, C. J. *Eco. Mod.* **172,** 109–139 (2004).
2. Fulton, E. A. *et al. Fish Fish.* **12,** 171–188 (2011).
3. Boyd, I. L. *Science* **337,** 306–307 (2012).
4. Mora, C., Tittensor, D. P., Adl, S., Simpson, G. B. & Worm, B. *PLoS Biol.* **9,** e1001127 (2011).
5. Brown, J. H., Gillooly, J. F., Allen, A. P., Savage, V. M. & West, G. B. *Ecology* **85,** 1771–1789 (2004).

Full author affiliations accompany this article online at go.nature.com/ob2a2p.


important to
the power of biomedical and matics,
achieving diversity in the biomedical research
workforce, and g a comp
and sustainable biomedical research
In June, these working group
recommendations, and at the yesterday
and this morning, we proposed how
implement them.
2011 2012
e overall success rate for research project grants (RPGs)
stayed the same compared to 2011. 18% 18%
The average size of RPGs increased. $449,644 $454,588
In 2012, there was an increase in the total amount of funding
that went to RPGs. $15,815,319,592 $15,92 ,746,065
NIH received more R01 grant applications. 28,656 29,515
Success rates for research using the mechanism
remained
The number eased. 5,264 5,340
NIH received more R21 grant applications. 13,
mechanism increased.

Two years of blogging the NIH

Sally Rockey, deputy director for extramural research at the US National Institutes of Health, reflects on the second anniversary of her precedent-setting blog.

ILLUSTRATION BY DAVID PARKINS

In 2010, a few of my staff members began pressuring me to start a blog. Although my friends and colleagues will attest that I'm not a shy or soft-spoken individual, I was, to put it mildly, reticent. Blogging was relatively new territory for government officials at the time. There were some outstanding examples of blogs by science-oriented government offices, such as the White House Office of Science and Technology Policy, which by then had been running for a couple of years (www.whitehouse.gov/administration/eop/ostp/blog). Jeremy Berg, former director of the National Institute of General Medical Sciences, had also successfully established a following on the blog Feedback Loop (https://loop.nigms.nih.gov), so I knew that it could be done well. But how would people respond to a National Institutes of Health (NIH) blog on extramural research policy? And how would it fit in to our work of supporting biomedical research?

My staff knew that I wanted transparency to be a centrepiece of my tenure at the NIH, and felt that the biomedical research community would find it refreshing to hear more from NIH management. Revealing how we do our business as we are developing policies and processes, and presenting data and information that shed light on the impact of our decisions, seemed like the right thing to do.

On 19 January 2011, I launched the Rock Talk blog (http://nexus.od.nih.gov/all/rock-talk). Two years later, I'm delighted that I did (even though I have taken heat for using too many exclamation marks!). As of 2 January 2013, I've written 107 blog posts, averaging about 40,000 page views a month. A steady stream of comments keeps me on my toes (see 'Top ten hits'). For the first time, the NIH was presenting data from all of its institutes and centres, and was sharing its analysis of funding trends on a forum where scientists could give their own opinions and reactions directly.

My first data-sharing post looked at trends in funding at various types of research institution — showing, for instance, that the majority of NIH funds go to medical schools, and that distributions to for-profit institutions have increased since the 1980s. Within two weeks, the post generated nearly 90 comments, about everything from the ethics of training non-US scientists to the importance of including more MD-trained investigators.

A SENSE OF COMMUNITY

We learned a lot about our constituents' interests and needs through the blog, and we have been able to highlight behind-the-scenes data, actively engage the community in policy-making and provide insight into our decisions. Without public input, effective and impactful policy cannot be created. My blogging experience has convinced me that using social-media platforms is one effective way for science-funding agencies

TOP TEN HITS

Ranked by number of individual page views, the ten most popular entries on the NIH Rock Talk blog generated varying numbers of comments.

Blog post title	Summary	Number of comments
1. Age Distribution of NIH Principal Investigators and Medical School Faculty	Compares the average age of NIH principal investigators (PIs) to that of medical-school faculty members, showing that more scientists are becoming PIs later in life and retiring later.	63
2. Paylines, Percentiles and Success Rates	Explains the relationship between the various scores applicants receive for grants, and how those correlate to success rates.	52
3. 2011 Success Rates, Applications, and Investigators	Statistics on funding for the 2011 fiscal year.	12
4. What's Behind the 2011 Success Rates?	Details on the 2011 success rates, and an explanation of how to find this information on RePORT.nih.gov.	46
5. How Do You Think We Should Manage Science in Fiscally Challenging Times?	Interactive data on the effect of potential policy changes, and an invitation for scientists to comment on how best to manage the NIH during times of austerity.	253
6. Does Your Degree Matter?	Funding rates for NIH applicants with PhDs, MDs or both.	19
7. The A2 Resubmission Policy Continues: A Closer Look at Recent Data	Data showing that preventing applicants from resubmitting proposals twice enables the NIH to award more proposals earlier in the application process, and doesn't penalize young investigators.	79
8. President's 2013 Budget and NIH Research Grants	Proposed 2013 budget and follow-up post to 'How Do You Think We Should Manage Science in Fiscally Challenging Times?'.	56
9. Postdoctoral Researchers — Facts, Trends, and Gaps	A closer look at postdoc support, scientists' salaries and time to first tenure-track job, using data from the US National Science Foundation's Survey of Earned Doctorates.	27
10. Piloting the $1.5M Special Review	Announcement of the decision to add an extra layer of review for well-funded applicants.	55

to successfully support research.

The blog is not the official vehicle for communicating policy changes. We have the NIH Guide for Grants and Contracts for that. And we have official channels for soliciting public feedback — through requests for information in the NIH Guide, for example. But the blog allows me to extend that conversation to people worldwide, many of whom I would not be able to reach in other ways. People are sometimes concerned that offering a dissenting opinion to NIH officials might affect their chances of getting funding. Although that is absolutely not the case, one advantage the blog does have over some other channels is that it allows people to remain anonymous if they wish.

Another attribute of the blog — which I think should be mirrored by most government blogs — is how we communicate current events in real time. For example, when Hurricane Sandy hit the US east coast in October 2012, we posted advice on what awardees should do if their labs were affected by the storm, and highlighted relevant resources and guidance for dealing with natural disasters.

I am often asked whether I really read and respond to comments in person. I do check them every day, on average; some I answer myself and some with the assistance of my blog team (made up of four staff who help regularly, plus five or so who contribute when needed). At times, the criticism of the NIH or even of me is harsh, and can feel personal. But I understand. In this time of tightening budgets, the topic of funding is very personal, and I understand the commenters' frustrations. The blog is moderated and we have had to screen out a very few comments that are off-topic, constitute a personal attack on another commenter or discuss the details of individual grants or applications (see the commenting policy at go.nature.com/ipwxqs).

The blog post that has received the most comments to date, 'How Do You Think We Should Manage Science in Fiscally Challenging Times?', was also one of the shortest. In fewer than 300 words, it asked scientists for their thoughts on how best to manage the NIH during times of austerity. This level of transparency into our thinking about future funding strategies was a novelty. The post also linked to interactive data charts that allowed readers to manipulate variables to see how their proposed changes to the size and numbers of grants we fund would affect success rates. We received more than 250 comments, including suggestions to take a closer look at how much funding was being given to individual scientists. We reviewed the suggestions and instituted our new Special Council Review policy to provide additional consideration of applications from principal investigators with existing NIH research project grants of US$1 million or more in direct costs (see J. M. Berg *Nature* **489,** 203; 2012).

"Conversation is an integral part of the policy-making process."

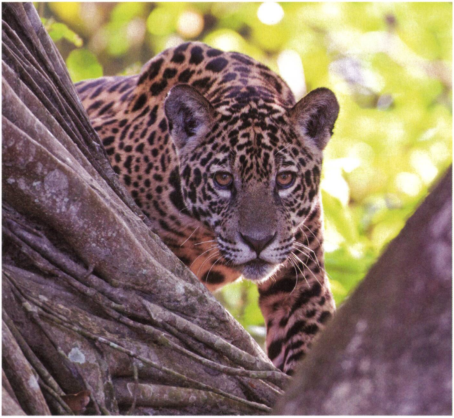
Blog posts containing data have clearly been the most popular (see table). Another common theme of my posts is debunking urban myths about NIH-supported research. Ideas for these posts come from blog readers' comments, grantee e-mails, discussions with attendees at research conferences and more. For example, in 'Myth Busting: Number of Grants per Investigator' we showed that the myth that many principal investigators continuously add to their bounty by piling on more and more NIH awards is untrue. The data we provided showed that the investigators who received the top 20% of funding in 2009 averaged only 2.2 grants each. (This example also illustrates how readers interact with the blog — they noticed that we had posted the wrong information in one of our graphs, which led to a follow-up post correcting the information and elaborating on the data we presented.)

Blogging is a lot of work, but it is also fun and enlightening. It would not be possible without the help of my Rock Talk blog and data teams, who provide statistical analyses, help me to decide on which topics to cover and moderate and respond to comments.

I encourage others in science policy to go into blogging with their eyes open, recognizing the level of effort required to care for and maintain a blog. Potential bloggers should be aware of the long-term commitment needed to determine which issues are blog-worthy and of interest to the community. Rock Talk has sometimes covered topics that reveal NIH idiosyncrasies. But I think that is what makes the blog real and helps us to connect with the community.

If you are not a reader of Rock Talk, I hope you will check it out. If you are, hopefully you know by now that I welcome your ideas, through either commenting on the blog or even starting a blog of your own. Conversation is an integral part of the policy-making process, and helps the biomedical research enterprise to keep on rockin'. ■

Sally Rockey *is deputy director of extramural research at the National Institutes of Health, Bethesda, Maryland 20892, USA.*
e-mail: sally.rockey@nih.gov



Jaguars were scarce even before they came under threat from humans.

MARK CARWARDINE/ARDEA.COM

CONSERVATION

The commonness of rarity

Stuart Pimm enjoys an exploration of what we can learn from rare species, if we care enough to keep them.

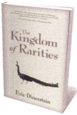
No writer of fiction would dare to invent the great naturalist Alfred Russel Wallace's adventures. In 1848, he funded his trip to the Amazon by collecting butterflies. Returning years later, his ship burned and sank mid-Atlantic; he lost his specimens, survived, and then set out for Asia. From Sarawak, Wallace nailed evolution's key consequences: "Every species has come into existence coincident both in space and time with a pre-existing closely allied species." He did it in so few words, he could have published it in *Nature*. But the scientific world largely ignored him.

NATURE.COM
For more about Alfred Russel Wallace, see:
go.nature.com/s9z8op

When Wallace later worked out what caused these laws — that is, these patterns in both time and space — he wrote to Charles Darwin for support. What islands in tropical Asia did for Wallace, the Galapagos had already done for Darwin. Familiar history follows. Less familiar is that Wallace and Darwin had discovered ecological rarity, the subject of Eric Dinerstein's engaging new book.

As Dinerstein explains in *The Kingdom of Rarities*, "many, many species on Earth are rare, but few people … are even aware of this". He gives us a kind of zoological travelogue, observing rare species across the planet and contemplating, as he does so, why rarity is profoundly important for our understanding of nature and our efforts to conserve it. Dinerstein — who is lead scientist of conservation science at wildlife charity WWF — starts by explaining that there are two kinds of rarity. Some species are widespread but sparse. Others are found only in very special places.

Large predators, such as jaguars, are an example of the sparse, widespread species. Dinerstein tracks these animals in South America and it soon becomes obvious why they are rare. Large-bodied species need a lot of food — and that demands large areas to support enough prey. Large fierce predators figure prominently in our imagination ("Lions and tigers and bears, Oh my!" says Dorothy in *The Wizard of Oz*). Dinerstein explains why Dorothy's fears are increasingly unfounded. We hunt these creatures to extinction — all too easy, given that they are rare to start with. This sets further problems in train: their herbivorous prey often become superabundant and decimate vegetation on which yet other species depend.



The Kingdom of Rarities
ERIC DINERSTEIN
Island: 2013. 312 pp. $29.95

Most of the book is, however, about rare species in special places, giving it a flavour of high adventure. We travel to a remote mountain top in the Indonesian province of Papua, accessible only by helicopter, through terrifying descents in thick mist. The species here are so rare that some are new to science. They live in isolation from "closely allied species" — and isolation from humans, too, as is clear from the abundance and tameness of large-bodied species in the area. There are species of marsupial tree-kangaroo, but no monkeys, leopards or tigers. As Wallace proposed, those species could walk to Bali and Borneo, but no farther, when sea levels were lower during glacial advances.

Dinerstein then takes us to mainland Southeast Asia, travelling with the first Western scientists to go there after the Vietnam War. Extensive bombing and the bush-meat trade make this an unlikely place for rare species to survive. Indeed, the rhinoceros (*Rhinoceros sondaicus*) is now extinct in Vietnam. In Hanoi, Dinerstein unexpectedly finds himself sitting next to General Vo Nguyen Giap, a "superb tactician" in the conflict with the United States. Now a conservationist, Giap says "the forest is our friend", adding that "it hid us and provided shelter". Dinerstein thinks that Vietnam has become "the last sanctuary for the region's forest-dwelling rarities": the forest hid them too.

STEVE GETTLE/MINDEN PICTURES/FLPA

Unexpectedly, Dinerstein then takes us to the United States, to Michigan, to visit the Kirtland's warbler (**pictured**), another rare species with a small range. Its existence suggests that the special places that host rare species can be anywhere, but then begs the question of where rare species are found most often. There are salient laws about diversity, beyond those that Wallace described.

First, most species have small ranges. Only a few species have large geographical ranges, but they are by far the most familiar to us. I've watched common kingfishers (*Alcedo atthis*) on British streams and on lakes in Japanese cities. To see other species in the same genus, however, I had to travel to Madagascar and Bali, and would need to travel to different islands in the Philippines to see two more.

Second, species with small ranges tend to be concentrated in special places — such as coastal Brazil or the eastern Himalayas. Most are tropical. Moreover, these areas are generally not where the greatest numbers of species are found in one place. So Dinerstein doesn't dwell on the Amazon: it has many species in any given location, but most have large ranges.

I sense that the Amazon disappointed Wallace too. It provided species, but not the unique ones his patrons craved. Darwin had no better luck in Patagonia, where the *Beagle* spent most of its famous voyage. When Wallace finally tripped over species with small, idiosyncratic ranges in the islands of Southeast Asia, and Darwin stopped briefly in the Galapagos on the way home, species rarity changed science forever.

Neither Darwin nor Wallace visited Hawaii, where Dinerstein's most poignant chapter is set. Different Hawaiian islands once had different "closely allied species". Dinerstein marvels at birds with strange names — the *'akiapola'au*, for example — and even stranger beaks, and at lovely lobelias with flowers that fit other strange beaks. He sees a few of these in the wild; but most are extinct, just specimens in museums. Those that remain hang on in low numbers in the face of numerous threats such as habitat loss. When species such as these disappear forever, they take into oblivion all that we might learn about rarity and all that it tells us about the natural world. ■

Stuart Pimm *is professor of conservation at the Nicholas School of the Environment, Duke University, Durham, North Carolina, USA, and author of* The World According to Pimm: a Scientist Audits the Earth.
e-mail: stuartpimm@me.com

STATISTICS

Number-crunching in the raw

David Vaux welcomes a primer on how statistics mesh with life, from salesmen's spin to political polls.

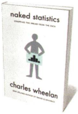
Ever wanted to encapsulate a quarterback's performance in a single number, wondered whether a bowling score was a descriptive statistic, or wished for a mathematical dissection of the Monty Hall box-choice problem? Then this is the book for you.

Like St Jude, Charles Wheelan has an interest in saving lost causes, and he has been invoking nudity to help in his quest. In an earlier book, *Naked Economics* (W. W. Norton, 2002), he attempted to bring some joy to the 'dismal science'. With *Naked Statistics: Stripping the Dread from the Data*, he faces another tough challenge: after all, statisticians have been described as people who are good with numbers but lack the personality to be accountants.

But this book is not intended for statisticians, because it doesn't go into the theory or mathematics in depth. Nor is there enough detail for clinicians, researchers or students. Instead, *Naked Statistics* targets the layperson, especially North American sports fanatics who want a deeper understanding of baseball, football, basketball, bowling, running and shooting statistics. But there is still plenty of thought-provoking material for the rest of us — even an amateur statistician like me — on stat-rich areas from politics, commerce and polling to the stock market.

There are, for instance, chapters explaining the spins used by salesmen, advertisers and politicians. One example is the blind taste test set up by a US beer company in the early 1980s, involving people loyal to another brand. Roughly half chose the tester's beer, but as Wheelan notes, the similar taste of mid-range US beers could mean that the result was down to the law of probabilities. Wheelan is excellent on political polls, margins of error, and the importance of questioning a truly representative sample of the population, which is why surveys of sexual practices and internet polls can generate spurious results.

There is sound advice on what Wheelan reveals as different types of modern gambling. He explains how in the stock market you cannot expect to do better than the market index; that the only insurance you should buy is when you can't afford to easily cover the loss yourself; and that lottery tickets are for chumps. I would have liked more on the historical role of gambling, and how it was used by Fermat and Pascal to establish the mathematical basis of probability and statistics. It would also have been welcome to have practical tips other than "Don't do it", such as the theory and practicality of counting cards at casinos, and how to compare and judge risk.



Naked Statistics: Stripping the Dread from the Data
CHARLES WHEELAN
W. W. Norton: 2013. 302 pp. $26.95

Naked Statistics also covers descriptive statistics (which allows you to describe how a population is distributed), inferential statistics (which can be used to compare populations or answer hypotheses), and correlation and regression analysis (which can be used to see whether data, such as diet and health, are related). Wheelan explains their uses and key shortcomings, but without formulae and problems to solve. The appendices to some of the chapters do include more mathematical detail, and will give you a deeper understanding, especially if you study them in conjunction with the superb, free, statistical videos on the Khan Academy website (http://www.khanacademy.org).

For a deeper explanation of statistics and a glimpse at a future in which we are not subject to the tyranny of p values — the most misunderstood statistical concept in research — I recommend Geoff Cumming's *Understanding the New Statistics* (Routledge, 2011).

And if you would like to know more about the areas Wheelan points to as hotbeds of use and abuse, take a look at *Proofiness* by Charles Seife (Viking, 2010) for the political angle, and *Bad Pharma* by Ben Goldacre (Fourth Estate, 2012), which looks at medical research and the pharmaceutical industry. These provide enough factual examples to get your blood boiling — without taking your clothes off. ■

David L. Vaux *is deputy director of the Walter and Eliza Hall Institute of Medical Research, Melbourne, Australia.*
e-mail: vaux@wehi.edu.au

BARBER SHOP
HAIR SALON
BARBER SHOP
Vukuzenzele Cash Store
Coca-Cola
HIV POSITIVE

The film *Fire In The Blood* focuses on the campaign for AIDS drugs in Africa.

SUNDANCE FILM FESTIVAL

FILM

Walking the edge at Sundance

Society and technology collide at this year's leading independent film festival, finds **Jascha Hoffman**.

Technology and imagination are cinema's warp and weft. So it is hardly surprising that the film industry has long been preoccupied with the leading edge of technological research — from *Metropolis* to *Transformers* — and with the ethical and social ambiguities that surround it. A range of movies tackle those grey areas at this year's Sundance Film Festival in Utah.

"We relate to the technology that we're offered," says John Nein, a senior film programmer at the festival. "We have to contend with the social and political forces around that technology."

The festival was founded in 1981 to help independent directors reach a wider audience. This year's science-and-technology-tinged offerings might convey less outright science than in years past, but they probe big issues — from the 'new' nuclear power to Google, WikiLeaks and the 1970s innovators who brought computers into our homes and our pockets.

➲ NATURE.COM
For the 2012 'science at Sundance' preview, see:
go.nature.com/6fcv94

A prominent pair of feature films deal with rivalry and uncertainty at the dawn of the digital era. Joshua Stern's *jOBS*, a biopic of Apple founder Steve Jobs, stars Ashton Kutcher as the young titan designing software and devices that will usher in the age of personal computing. *Computer Chess*, a comedy-drama about programmers teaching computers to play chess in the 1980s, underlines the programmers' anxiety to prove that computers could hold their own against humans. That remains a live issue as we begin to carry on conversations with the computers in our phones and cars. "The big philosophical questions may seem charmingly nostalgic, but lord knows we haven't answered them yet," director Andrew Bujalski wrote when raising a modest budget for the film online.

Two documentaries tackle controversial efforts to spread information on the Internet — and the resulting backlash. Ben Lewis's *Google and the World Brain* charts the Internet giant's decade-long effort to scan the world's books into a giant digital archive, prompting legal action from authors and librarians. And from prolific documentary director Alex Gibney comes *We Steal Secrets*, which traces the evolution of WikiLeaks from a site founded in Sweden in 2006 by Australian former hacker Julian Assange to an international movement that aims to make secrecy obsolete with actions such as the release of 250,000 US diplomatic cables in 2010 and 2011.

Medicine features heavily in another brace of documentaries, both of which show researchers and advocates racing against the clock to devise and disseminate cures. Sean and Andrea Fine's *Life According to Sam* traces the efforts of husband-and-wife US doctors Scott Berns and Leslie Gordon to understand the rare premature-ageing disease progeria, after their son was diagnosed at age two. Their work over the ensuing decade has led to gene discovery, preliminary drug trials and potential insights into the ageing process. Africa at the turn of the millennium is the focus of Dylan Gray's *Fire in the Blood*, which follows the public-health activists, including South African Zackie Achmat, who pushed governments and drug companies to make AIDS drugs available to Africans.

Even the films with an environmental theme have a technological edge. *Pandora's Promise*, directed by Robert Stone, is a documentary about environmentalists such as the British journalist Mark Lynas, who, after opposing nuclear power, now support the use of new technologies such as small modular nuclear reactors to reduce reliance on fossil fuels. The eco-thriller *The East*, directed by Zal Batmanglij and starring Brit Marling and Ellen Page, features an activist group that attacks corporations using their own products, such as harmful chemicals.

Some of the festival's more conceptual works use technology to encourage viewers to empathize with other creatures. *Coral: Rekindling Venus*, is a full-dome planetarium show by Australian artist Lynette Wallworth. Deep-sea photography and music by Antony Hegarty (of the New-York-based group Antony and the Johnsons) among others highlight the threats to coral reefs. The short film *Primate Cinema: Apes as Family*, by artist Rachel Mayeri, shows the reactions of real chimpanzees to a television programme acted by humans in chimp costumes, inviting us to ponder how our fellow primates see us.

A panel on science at the movies, organized by the Alfred P. Sloan Foundation's Science and Film programme, promises a frank discussion of "the best and worst of cinema science" while honouring the "creative, imaginative, speculative and even adventurous" spirit of scientists and film-makers. Panellists, including director Darren Aronofsky and theoretical physicist Lisa Randall, will also select a winner for this year's Sloan Prize for films about technology and science. ■

Jascha Hoffman *writes for* Nature *and* The New York Times. *He lives in San Francisco.*
e-mail: jascha@jaschahoffman.com

nature.com webcasts

nature | methods presents a webcast on

Fluorescent proteins and sensors:
A practical discussion

Fluorescent proteins have become invaluable tools for fluorescent microscopy in the life sciences but researchers still have many practical questions about using them. Three experts in the development and use of fluorescent proteins and sensors discuss these challenges and provide practical advice to users.

Now live on demand

Register for FREE:
www.nature.com/webcasts/fluorescent_proteins

Sponsored by: CHROMA

nature publishing group npg

Correspondence

Rat reality show blurs quality control

The governance of science, including the whole system of quality assurance, depends on specialist access to resources and publication. A new and radical engagement of the public in reality and crowd-sourced science is calling this principle into question.

Following the furore surrounding cancer claims for herbicide-resistant transgenic maize (*Nature* **489,** 484; 2012; and *Nature* **492,** 12; 2012), Russian scientists intend to run a 'rat reality show' that will be broadcast over the Internet around the clock. This year-long, real-time feeding experiment aims to test the safety of genetically modified food products (see go.nature.com/qkvzqe). Actually witnessing the experiment, the team suggests, will allow the public to draw their own conclusions. If the scientists are unable to fund the show from orthodox sources, they plan to crowd-source public funding.

It would be easy to dismiss such 'reality' experiments as a stunt — as frivolity leading to demagoguery. But social media are increasingly influencing mainstream scientific communication and could stimulate a spread in reality science, blurring the demarcations on which the legitimacy and quality assurance of science traditionally depend.

Jerome R. Ravetz, Peter Healey, Steve Rayner *Institute for Science, Innovation and Society, University of Oxford, UK.*
steve.rayner@insis.ox.ac.uk

TOXIC
SAFE
TOXIC
SAFE
SAFE
TOXIC
SAFE
TOXIC
SAFE
TOXIC
S. Harris

Cancer goal: deadline will not erode trust

We at the US National Breast Cancer Coalition (NBCC) disagree with your assessment of our 2020 goal of knowing how to prevent breast cancer and its associated deaths (*Nature* **491,** 637; 2012). Working towards that goal will not erode public trust, as you suggest.

Nearly 500,000 women worldwide died of breast cancer last year, despite billions of dollars being invested in research. Many scientists believe that current funding systems favour 'safe' research over bold new ideas. As a result, progress is incremental, leading to slightly better treatments, surgical interventions and radiation regimes. These may provide some benefit, but bring no end to the disease itself.

Scientists and the NBCC need to work together to reorder priorities and change the conversation and culture of science. Trust is not lost when advocates call for a deadline and provide a blueprint for meeting it. But it is eroded when scientific infrastructure is unaccountable to the people intended to benefit from its output; when there is not enough emphasis on translating research discoveries to the clinic; and when published results cannot be replicated and marginal advances are over-hyped. Meanwhile, more and more people lose their lives.

Take a calculated risk with us. Let's reach for what might in fact be possible.

Frances Visco *National Breast Cancer Coalition, Washington DC, USA.*
fmvisco@breastcancerdeadline2020.org

Cancer goal: vaccine is cause for optimism

As the originator of the project to develop a preventive breast-cancer vaccine, which is supported by the US National Breast Cancer Coalition's (NBCC) 2020 deadline, I question your dismissal of this new approach (*Nature* **491,** 637; 2012).

You base your pessimism on the genomic complexity of tumours and the length of time that clinical trials would take to test such a vaccine. But the genomic complexity apparent in mature tumours is not relevant for developing a preventive vaccine: the immune system needs only to be preactivated with antigens presented by the nascent tumour. Unstable RNA processing in a tumour is likely to be a richer source of these antigens than genomic mutations.

It could indeed take a decade or more to validate the treatment using today's clinical standards. But trials could be made shorter and less expensive if, for example, there were definitive ways to detect very early tumours.

You say that "discovery does not answer to deadlines", but accomplishment can. Given the alternatives, we should embrace this effort.

Stephen Albert Johnston *Arizona State University, Tempe, USA.*
stephen.johnston@asu.edu

Tie carbon emissions to consumers

Global carbon dioxide emissions soared from 22.7 billion tonnes in 1990 to 33.9 billion tonnes last year, despite 20 years of attempted mitigation (*Nature* **491,** 656–658; 2012). The sizeable economic gaps between nations are largely responsible for the international deadlock in climate negotiations. A radically new approach is needed.

One solution would be to allocate common but differentiated responsibility for mitigating emissions to individual consumption activities, rather than to countries. Profligate consumers from both developed and emerging countries are the worst offenders for generating non-essential emissions. They should be held accountable for those emissions, irrespective of where they are generated.

Emissions produced after reallocating responsibility in this way could be quantified for nations, lifestyles or even every product consumed. This could be achieved by integrating the 'top-down' consumption-based accounting methods that are used to determine national total consumption emissions with 'bottom-up' carbon-footprint calculations based on analysis of products' life cycles (G. P. Peters *Curr. Opin. Environ. Sustain.* **2,** 245–250; 2010). A living standard, together with a per capita emission quota, would be defined so that people worldwide could meet their basic living requirements without mitigation costs.

Such standardized measures would allow new cap-and-trade policies and carbon-taxing mechanisms to run smoothly and effectively across different consumption groups at a global scale (see M. Grubb *Nature* **491,** 666–667; 2012).

Zhu Liu, Fengming Xi *Institute of Applied Ecology, Chinese Academy of Sciences, Shenyang, China.*
liuzhu@iae.ac.cn
Dabo Guan *University of Leeds, Leeds, UK.*

Sexism: measure journal objectivity

Your initiative to address the issue of sexism in publishing is laudable (*Nature* **491,** 495; 2012). But scientific editors may be biased in other ways that influence the publication process — and they are not in the best position to recognize and correct their own biases.

An independent, external assessment body that regularly evaluates editorial practices across scientific journals might be the answer. This would promote transparency and reassure authors that their work is being dealt with fairly. An 'objectivity factor' resulting from such an assessment could become a key metric of journal performance, alongside its impact factor.

Boyan K. Garvalov *University of Giessen, Germany.*
boyan.garvalov@patho.med.uni-giessen.de

Sexism: control experimental bias

Implicit bias against women (*Nature* **491,** 495; 2012) has a record of influencing the design of experiments and collection of data in the life and mind sciences.

Over the past four decades, feminist scientists, historians and philosophers of science have presented case study after case study showing how sexist bias can distort scientific results. Examples include work by Rebecca Jordan-Young on brain-organization theory, Anne Fausto-Sterling on the biological study of sex differences and Anelis Kaiser on neuroimaging.

In the interest of greater accuracy, controls for sexist bias need to be more rigorous.

Letitia Meynell *Dalhousie University, Halifax, Nova Scotia, Canada.*
letitia.meynell@dal.ca

Sexism: conferences should seek a balance

Conference committees and symposia chairs should follow *Nature*'s lead and be more open and proactive about gender balance (*Nature* **491,** 495; 2012).

Publishing an online declaration of gender-balance policy would help. Data from each of the past, say, three meetings could also be posted online to indicate the percentage of women who were registrants, invited plenary and keynote speakers, session speakers, programme committee members, or session chairs. If there is a significant gap between the percentage of women attending the conference and the overall percentage of women speakers or committee members, then a policy overhaul may be called for.

Jennifer L. Martin *University of Queensland, Brisbane, Australia.*
j.martin@imb.uq.edu.au

More trials needed to assess sleeping pills

Sam Fleishman suggests that the controlled use of sleeping pills helps to counter the life-disrupting consequences of insomnia (*Nature* **491,** 527; 2012). But, to our knowledge, there is no convincing evidence that sleep medication can reactivate the health-restoring functions of sleep.

Good-quality sleep improves cognitive performance, vigilance, memory and mood. Poor sleep is associated with potentially damaging physiological effects such as inflammation (M. R. Irwin *et al. Brain Behav. Immun.* **24,** 54–57; 2010), compromised immune-cell activity (E. Fondell *et al. Brain Behav. Immun.* **25,** 1367–1375; 2011) and telomere shortening (A. A. Prather *et al. J. Aging Res.* **2011,** 721390; 2011).

However, the regular use of sleeping pills has been linked with increased mortality and morbidity from infection, depression or cancer (see, for example, D. F. Kripke *et al. BMJ Open* **2,** e000850; 2012).

Large randomizing trials are needed to evaluate the benefits and risks of sleeping pills so that they can be prescribed more responsibly.

Maximilian Moser *Medical University of Graz, Austria*
max.moser@medunigraz.at
Daniel F. Kripke *University of California San Diego, California, USA.*

Toe-clipping vital to amphibian research

Keeping a record of the global decline of amphibian populations depends on the identification and marking of individuals in the field, commonly by toe-clipping. This work is under threat: the Brazilian federal government and non-governmental organizations want to prohibit toe-clipping without scientific justification.

Toe-clipping is a simple, cost-effective, minimally invasive marking technique that has been in use for decades in herpetological research. The Brazilian federal agency that enforces environmental policies, IBAMA, is claiming that the practice is a form of mutilation and should be a criminal offence under federal law (see go.nature.com/qkij7l; in Portuguese). The Brazilian Herpetological Society has protested against this anthropocentric position (see go.nature.com/thpwdg; in Portuguese) on the grounds that it would set back efforts to understand amphibian decline.

Brazil is home to the greatest amphibian diversity on Earth, but knowledge of population dynamics is scant and will remain so without access to reliable marking methods. We acknowledge that toe-clipping is not an ideal solution but, when properly implemented, it has minimal effects on amphibian survival and behaviour. Without this technique, it will be harder to obtain crucial information that could prevent amphibian species from becoming extinct (W. C. Funk *et al. Nature* **433,** 193; 2005).

Décio T. Corrêa* *State University of Campinas, São Paulo, Brazil.*
decio.tadeu@gmail.com
*On behalf of 5 co-signatories (see go.nature.com/5sgsqo for full list).

Missing data mean holes in tree of life

As part of the Open Tree of Life project (http://opentreeoflife.org), we surveyed publications covering all domains of life and found that most phylogenetic trees and nucleotide alignments from the past two decades have been irrevocably lost.

Of 6,193 papers we surveyed in more than 100 peer-reviewed journals, only 17% present accessible trees and alignments (used to infer relatedness). Contacting lead authors to procure data sets was only 19% successful. DNA sequences were deposited in GenBank for almost all these studies, but it is the actual character alignments that are pivotal for reproducing phylogenetic analyses. We estimate that more than 64% of existing alignments or trees are permanently lost.

This problem will increasingly hinder phylogenetic inference as the use of whole-genome data sets becomes common. Journals need to reinforce a policy of online data deposition, either as supplementary material or in repositories such as TreeBASE (http://treebase.org) or Dryad (http://datadryad.org) — including for data sets based on previously published sequences. Ecologists, evolutionary biologists and others will then have access to rigorous phylogenetics for testing their hypotheses.

Bryan T. Drew* *University of Florida, Gainesville, USA.*
bdrew@ufl.edu
*On behalf of 8 co-signatories (see go.nature.com/gzqcr9 for full list).

CORRECTION

The Outlook article 'Genetics: Searching for answers' (*Nature* **491** (suppl. 7422), S4–S6; 2012) incorrectly stated that the Simons Simplex Collection consists of 3,000 blood samples taken from more than 700 people. In fact, it contains 13,000 blood samples taken from 3,000 people.

Rita Levi-Montalcini

(1909–2012)

Nobel prizewinning neurobiologist and eminent advocate for science.

It is unusual to dominate a field so extensively as to be synonymous with it. Such was the stature of Rita Levi-Montalcini, who shared a Nobel prize for her discovery of nerve growth factor. Had her extraordinary life been scripted by Hollywood, it would have been rejected as too improbable.

Levi-Montalcini died on 30 December 2012, aged 103. She was born in 1909 to a well-to-do Italian Jewish family in pre-Fascist Turin. Rita Levi (she added her mother's family name as an adult) began her medical studies at the University of Turin in 1930, after convincing her autocratic father that this was an appropriate choice for a young woman. Her first mentor, Giuseppe Levi, kindled her interest in the nervous system and remained a colleague for many years.

In 1938, Levi-Montalcini was forced out of the university by Fascist race laws. This began a period of hardship, despair and perseverance. She carried out research on embryological neurodevelopment in her home 'laboratory', first in Turin, and later in Florence after her family was forced to flee. She did experiments using crude, hand-made dissection tools and fertilized chicken eggs, later eaten for dinner.

During this period, she read an article by Viktor Hamburger, an embryologist at Washington University in St. Louis, Missouri. He wrote that when the growing limbs of chick embryos were cut off, the resultant atrophy in the neuronal cell clusters intended to innervate them was due to a loss of an 'inductive factor' from the absent limbs. Hamburger suggested that this factor was necessary for the growth and differentiation of the neural precursor cells.

Levi-Montalcini repeated this experiment with Levi, using the silver-staining technique he had taught her. At first, the neurons continued to grow, dying only when they could not reach their surgically eliminated target. Levi-Montalcini concluded that the neuronal death resulted from the absence of a growth-promoting substance — rather than an inductive factor — released by the target of the growing neurons. After the war, Hamburger invited her to his laboratory in the United States to resolve their differing views.

As with most great discoveries, the identification of nerve growth factor (NGF) depended on serendipity. First, a student of Hamburger's, Elmer Bueker, observed that a fast-growing sarcoma became highly innervated when transplanted into a chick embryo — providing a testable source of NGF. Then Stanley Cohen, in the same lab, discovered that snake venom and the submandibular glands of adult male mice were even richer sources (Cohen joined Levi-Montalcini in the biochemical characterization of NGF).

HERB WEITMAN/BECKER MEDICAL LIBR., WUSTL.

In 1952, in Rio de Janeiro, Brazil, Levi-Montalcini captured the first image of the production of nerve fibres growing under the influence of NGF from chick-embryo neurons that had been excised and placed in culture in the laboratory of Hertha Meyer, a friend from student days and an expert in the required techniques. With the aid of two tumour-bearing mice, smuggled into Brazil in Levi-Montalcini's handbag, the women developed the first 'halo' assay of NGF. This dramatic bit of scientific art confirmed the existence of the target-derived growth-promoting substance she had predicted; it was one of her favourites.

The discovery revolutionized neurobiology and endocrinology, and set the stage for cell-signalling research. But, at first, NGF was not universally accepted as significant. Levi-Montalcini had to deal with many a naysayer until the early 1970s, when evidence for the importance of NGF (and many similar factors) became overwhelming. She was also not a biochemist, and when Cohen left St Louis in 1959, she had to leave the molecular characterization of NGF to others. She prided herself on her intuition, but rebelled against the idea that NGF was a hormone-like substance and that it was representative of a larger group of growth factors with broad specificities. That said, she, with colleagues in Rome, did subsequently contribute important observations about its expanded role, particularly in the immune system and in ophthalmology.

Levi-Montalcini remained on the faculty of Washington University for 30 years. Missing her twin sister, Paola, she began a gradual return to Italy during the 1960s. In 1969, she became the director of the new Italian Research Council's Laboratory of Cell Biology in Rome, which grew to be one of the largest research centres in the country.

In 1986, with Cohen, she was awarded the Nobel Prize in Physiology or Medicine — becoming the first Italian woman so honoured. This launched her as the omnipresent advocate for Italian science. Over the next quarter of a century, she participated in television debates, wrote popular-science books and provided advice to the government and other agencies. In 2001, she became a senator for life in the Italian government, influencing key policy questions.

In 1992, she established the Rita Levi-Montalcini Foundation, which supports women from developing countries — particularly Africa — in finding careers in science. She also founded the European Brain Research Institute in Rome in 2002.

Rita has been described as ambitious, autocratic, generous, possessive, aristocratic, demanding, persevering, insightful and totally dedicated to her work. All are accurate. She had disagreements with many scientists who worked on NGF. However, as the importance of her discoveries became increasingly appreciated, she mellowed in her outlook (if not her drive) and seemed to accept the mantle of matriarch that was truly her due.

In April 1986, I co-organized the first meeting on NGF, in Monterey, California, to coincide with her 77th birthday (complete with cake and candles). Six months later, when she and Cohen were awarded the Nobel prize, I sent congratulations. She replied that while it was a wonderful recognition of her life's work, she enjoyed her party more. This diminutive woman was a giant in science. ■

Ralph A. Bradshaw *is professor emeritus in the Department of Physiology and Biophysics at the University of California, Irvine. He collaborated with Levi-Montalcini on the characterization of NGF.*
e-mail: rablab@uci.edu

RESEARCH

NEW ONLINE

Papers published this week at Nature.com/nature

NEWS & VIEWS Microscopy that defines defects in metal–organic frameworks **p.313**

LETTER Teasing out the factors that contributed to the recent East African drought **p.389**

LETTER A surprising genetic link to the burrowing architecture of a rodent **p.402**

PUBLISHED 13 JANUARY 2013

IMMUNOLOGY

'Good' microbes send right signals

An important question in the field of intestinal immunology is how the immune system can mount protective immune responses against pathogens while avoiding such responses against useful or commensal organisms. Gretchen Diehl *et al.* show that commensal bacteria inhibit the induction of mucosal immune responses by providing signals that limit the trafficking of phagocytes within the lamina propria that have high levels of the chemokine receptor CX$_3$CR1 to the mesenteric lymph nodes. Therapeutic modulation of these cells may attenuate intestinal inflammation or enhance priming for mucosal vaccines.

Microbiota restricts trafficking of bacteria to mesenteric lymph nodes by CX$_3$CR1^{hi} cells
G E Diehl *et al.*
doi:10.1038/nature11809

PALAEONTOLOGY

Role reversal of tetrapod vertebrae

A backbone of interlocking vertebrae is a prerequisite for a tetrapod — a vertebrate that lives on land. However, skeletal details are very hard to make out in fossils of the very earliest tetrapods because the specimens are often swathed in rocky matrix. Application of the latest synchrotron microtomography technology to the three best-known early tetrapods — *Ichthyostega*, *Acanthostega* and *Pederpes* — has now overcome that inconvenience. Early tetrapods were presumed to have had 'rhachitomous' vertebrae, with a neural arch and spine above, a single intercentrum front and ventral and paired pleurocentra to the rear and ventral. But the microtomography scans reveal a 'reverse' rhachitomous design, effectively rewriting the 'textbook' view of vertebral column evolution in the first limbed vertebrates.

Vertebral architecture in the earliest stem tetrapods
S E Pierce *et al.*
doi:10.1038/nature11825

CANCER

Reciprocal regulation of p53 and malic enzymes modulates metabolism and senescence
P Jiang *et al.*
doi:10.1038/nature11776

STRUCTURAL BIOLOGY

Rotation mechanism of *Enterococcus hirae* V$_1$-ATPase based on asymmetric crystal structures
S Arai *et al.*
doi:10.1038/nature11778

Structural basis for viral 5′-PPP-RNA recognition by human IFIT proteins
Y M Abbas *et al.*
doi:10.1038/nature11783

DEVELOPMENTAL BIOLOGY

Mitotic cell rounding accelerates epithelial invagination
T Kondo & S Hayashi
doi:10.1038/nature11792

PUBLISHED 16 JANUARY 2013

ECOLOGY

Pushing biofuel into the margins

Agricultural land that could be used for much-needed food crops is being diverted towards production of grain-based ethanol in both Europe and the United States, in part as a response to government biofuels legislation. Cellulosic crops could be grown instead on land that is poorly suited to field crops. An evaluation of the potential of such 'marginal' lands in the midwestern United States to produce biofuel while mitigating direct greenhouse-gas emissions now finds that a significant amount of biofuel energy could be produced without the initial carbon debt and indirect land-use costs associated with food-based biofuels.

Sustainable bioenergy production from marginal lands in the US Midwest
I Gelfand *et al.*
doi:10.1038/nature11811

Biofuel production on the margins
K Butterbach-Bahl & R Kiese
doi:10.1038/nature11853

NEUROSCIENCE

Light vital to developing eye

As the retinal vasculature forms in the developing eye, the blood vessels supplying the transparent hyaloid membrane surrounding the vitreous body regress. This process is important clinically because excessive vascular growth is a major cause of blindness in premature infants. Surprisingly, Richard Lang and colleagues have now found that light is involved in this major change in tissue architecture by stimulating regression of developing hyaloid vasculature in mice, acting through the photoreceptor melanopsin. In the absence of either light or melanopsin, the vascular endothelial growth factor A is up-regulated and abnormal retinal angiogenesis results. Whether the same processes are involved in human eye development remains to be determined.

A direct and melanopsin-dependent fetal light response regulates mouse eye development
S Rao *et al.*
doi:10.1038/nature11823

EVOLUTIONARY GENETICS

Y-like social chromosome

Fire ants (*Solenopsis invicta*) live in two very different types of colony: in some colonies, the workers tolerate multiple queens; in others, they tolerate just one. This divergence is in the genes, under the control of a single Mendelian factor. A large-scale analysis of the genomic region involved shows that the various aspects of this polymorphism are governed by a non-recombining supergene that occupies half a chromosome. The operation of this region is remarkably similar to that of sex chromosomes. The findings show how limiting local genetic rearrangements can maintain divergent social behaviours that involve many genes acting together.

A Y-like social chromosome causes alternative colony organization in fire ants
J Wang *et al.*
doi:10.1038/nature11832

Forum: A social rearrangement
A F G Bourke; J E Mank
doi:10.1038/nature11854

IMMUNOLOGY

A T-bet gradient controls the fate and function of CCR6$^-$RORγt$^+$ innate lymphoid cells
C S N Klose *et al.*
doi:10.1038/nature11813

NEWS & VIEWS

EVOLUTION

A history of give and take

Fossils of microorganisms from deep-sea sediment cores show that environmental change correlates closely with extinction but not with speciation, producing a nuanced view of the drivers of evolution. See Letter p.398

STEVEN M. HOLLAND

Whether the diversification of life is driven by interactions among species or by physical and chemical forces is one of the most fundamental questions for understanding the biological history of Earth. On page 398 of this issue, Peters *et al.*[1] make a convincing case that the number of species of a major group of single-celled marine organisms is closely linked to historical changes in the physical and chemical properties of the ocean and atmosphere. Extinctions seem also to have mostly resulted from these environmental factors, but species origination does not, suggesting that speciation is primarily controlled by organismal interactions*.

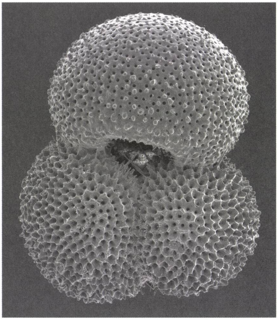


Figure 1 | Layers of life. This scanning electron micrograph shows the fossilized shell of the planktic foraminifer *Globigerinoides ruber*. Foraminifera are a class of amoeboid protist that is widespread throughout marine environments. Peters and colleagues[1] measured the diversity of fossilized foraminifera in deep-sea cores from the Atlantic Ocean, which also record changes in the physical and chemical environment.

DENNIS KUNKEL MICROSCOPY/PHOTOTAKE

Peters and colleagues studied the 170-million-year evolutionary history of planktic foraminifera, which are free-floating single-celled organisms, in 73 deep-sea sediment cores from the Atlantic Ocean. Foraminifera are excellent for studies such as this because they are extraordinarily abundant, are geographically widespread and have shells made of calcium carbonate, which readily preserve as fossils (Fig. 1). And the Atlantic Ocean is an ideal setting, because decades of sea-floor drilling have generated a richly documented stratigraphic record, which shows spatiotemporal changes in sediment that result from tectonic or climate-induced changes.

The authors found a strong correlation between the evolutionary history of planktic foraminifera and the waxing and waning of sediment deposition in the Atlantic. Normally, a correlation between the extent of sedimentation and the diversity of organisms would be interpreted as a sampling bias: when sedimentation is extensive, there are more opportunities for fossilization, so the apparent diversity of organisms increases[2,3]. Peters *et al.* convincingly refute the possibility of sampling bias. First, they argue that if sampling causes these patterns, it should affect species originations and extinctions equally. It doesn't. The geographical extent of the sedimentary rock record is substantially better correlated with extinction than it is with origination, a relationship that Peters has shown previously for other groups of marine invertebrates[4]. Second, if sampling is the dominant control, then the diversity of foraminifera should increase towards the present day as the sedimentary record becomes more extensive, reflecting the progressive widening of the Atlantic Ocean. Again, it doesn't. Third, the authors show that, on average, only seven locations are needed to capture all the foraminifera species existing at any time. This is far fewer than the dozens of cores that are typically needed to preserve sediment of any given age.

An alternative explanation for a correlation between the history of planktic foraminifera and the extent of deep-sea sedimentation — and one that fits Peters and colleagues' findings well — is the common-cause hypothesis, which states that both diversity and sedimentation are responses to environmental changes[4]. An example of common cause would be the global cooling during the Eocene–Oligocene transition (around 34 million years ago), which brought about the extinction of warm-water species, the geographic expansion of cool-water species, and increased preservation of deep-sea carbonate sediments. The finding of a strong coupling between deep-sea sedimentation and extinctions suggests that extinction is driven primarily by environmental perturbations. However, the lack of a link between sedimentation and species origination suggests that origination is governed more by species interactions.

This intriguing conclusion raises many questions that other palaeobiologists will surely address. One set of questions relates to macroevolutionary processes — those that occur at or above the level of species. For example, why do environmental perturbations drive extinction but apparently have little effect on origination? Might this indicate that origination takes longer than extinction, which would weaken its apparent correlation with the rock record? And what are the specific mechanisms that link environmental change and extinction?

Another series of questions arises from departures from the correlations that the authors observed in certain periods. For example, records of a meteorite impact at the end of the Cretaceous, around 65 million years ago, clearly explain why extinctions of foraminifera are unusually high at this time. But why is extinction so anomalously low in the time immediately following? Although origination and the sedimentary record are uncorrelated, there is much structure in the relationship of the two. For instance, why do all the largest decreases in origination rate occur at times when the stratigraphic record shows little or no change in sedimentation? Moreover, why are origination rates often unusually high during times in which deposition is not commencing in new areas?

The approach that Peters and colleagues take, of comparing evolutionary metrics with metrics of the stratigraphic record, is called macrostratigraphy[5]. This method has been instrumental in moving palaeobiology away from a preoccupation with sampling biases and towards seeking mechanisms of diversity change. The evidence in foraminifera for the common-cause hypothesis adds to previous evidence from marine invertebrates, but raises the question of how widespread these macroevolutionary relationships are. The common-cause hypothesis is one of the most interesting developments in macroevolution, and advances in stratigraphy and analytical methods will continue to have an important role in testing it[6]. ■

---

*This article and the paper under discussion[1] were published online on 9 January 2013.

Steven M. Holland *is in the Department of Geology, University of Georgia, Athens, Georgia 30602-2501, USA.*
e-mail: stratum@uga.edu

1. Peters, S. E., Kelly, D. C. & Fraass, A. J. *Nature* **493,** 398–401 (2013).
2. Raup, D. *Paleobiology* **2,** 289–297 (1976).
3. Smith, A. B. *Phil. Trans. R. Soc. B* **356,** 351–367 (2001).
4. Peters, S. E. *Proc. Natl Acad. Sci. USA* **102,** 12326–12331 (2005).
5. Peters, S. *J. Geol.* **114,** 391–412 (2006).
6. Patzkowsky, M. E. & Holland, S. M. *Stratigraphic Paleobiology* (Univ. Chicago Press, 2012).

EARTH SCIENCE

Plumbing nickel from the core

Magmas that have erupted at Earth's surface reveal a potential new mantle source. This source, which is rich in nickel and has a primordial helium isotopic content, may have originated at great depth in the mantle. SEE LETTER P.393

MICHAEL J. WALTER

Melting of source rocks in the mantle generates magmas that erupt at Earth's surface. Identifying chemically unique mantle source rocks is a big deal because they can reveal much about Earth's internal structure, which in turn informs our understanding of plate tectonics and planetary evolution. By far the best-defined mantle source is the part of the mantle that is extensively tapped by magmas called MORBs, which erupt at plate boundaries along mid-ocean ridges. Two basic properties of the source rock for these MORBs are that it contains about 2,000 parts per million of nickel and that its helium isotopic composition is a result of radioactive decay and is thus not primordial[1,2]. The MORB source rock defines the 'normal' upper mantle, and possibly dominates the mantle as a whole.

In this issue, Herzberg and co-workers[3] investigate the chemistry of 'primary' mantle-derived magmas from several locations away from plate boundaries*. (Primary magmas are those obtained directly by melting of the mantle, and which have not undergone any change in chemical composition.) The authors identify a new mantle source containing about 20% more nickel than that of normal upper-mantle rock. This source also seems to have a primordial helium isotopic composition. Herzberg *et al.* identify the location of this reservoir as the core–mantle boundary — a primordial source that is sometimes piped to the surface through mantle plumes.

Primary magma from the mantle can be used to probe the composition of the mantle source rock. But it is one of nature's foibles that primary magmas almost never make it to the surface because, as they rise and cool, they crystallize and so differentiate. This means that petrologists must devise models that can eliminate the effects of crystallization. Olivine is the first mineral to crystallize from most mantle melts. Fortunately, the co-variation of olivine's nickel content and 'magnesium number' (the ratio of the molar concentration of magnesium to magnesium plus iron) means that the amount of crystallization a magma experiences can be tracked, and this provides a way of estimating the composition of the primary mantle melt. Once the primary melt has been reconstructed, the composition of the mantle source rock can also be deduced. It is this kind of modelling that underpins Herzberg and colleagues' identification of a unique, nickel-rich mantle source.

However, accounting for the effects of crystallization is a tricky business, because the way in which nickel partitions between olivine and the melt can depend on the pressure, temperature and composition of the differentiating magma. Several competing empirical models for nickel partitioning are available, some of which involve significant temperature dependence[4,5]. The nickel partitioning model used by Herzberg *et al.* does not include a temperature term, and so if considerable temperature dependence does exist, then the calculated differences that the authors find in the nickel contents of primary magmas might simply be a temperature effect and not a reflection of a distinct mantle source[5]. Herzberg *et al.* are well aware of this pitfall, and provide cogent arguments to explain why temperature variation is an unlikely reason for the differences in nickel content that they have calculated in the primary melts. Nevertheless, among many aficionados, this temperature effect and its consequences remain an issue of active debate.

Another way in which a high nickel content might be found in olivine crystallizing from a primary magma would be for olivine not to have been present in the mantle source rock at all[6]. Currently, this is a leading explanation for the unusually high nickel abundance found in olivine from Hawaii and other magma locales lying within plates rather than at plate boundaries[7], although here again, temperature effects have been invoked as an alternative explanation[5]. In this 'olivine-free' source model, partial melts extracted from recycled oceanic crust react with the normal mantle and dissolve away the olivine. Herzberg *et al.* contend that this process does not work for their magmas because chemical diagnostics involving the iron, manganese and calcium content of olivine simply do not fit the model. If temperature effects and olivine removal are not the culprits, Herzberg and colleagues are effectively left with only one alternative — a new high-nickel-content source rock.

So where does the excess nickel come from? Herzberg *et al.* propose that it originates from the most nickel-rich reservoir known — Earth's core. It has been conjectured that, early in Earth's history, a primordial basal magma ocean existed at the core–mantle boundary[8], presumably in chemical equilibrium with the iron- and nickel-rich metallic liquid core. A solidified or partially molten vestige of this basal magma may have survived to the present, existing as a unique mantle reservoir.

Nickel is a 'siderophile', which means that it partitions strongly into the core relative to the mantle. But partitioning is known to be pressure dependent, with increasing pressure increasingly favouring the mantle[9]. Herzberg and colleagues surmise that, at the pressure that exists at the core–mantle boundary (about 130 gigapascals), the mantle should be enriched in nickel, whereas the rest of the mantle would have equilibrated with the metal at much lower pressures during core formation (about 50 GPa)[9]. This ancient remnant of basal magma would also retain a primordial helium isotopic composition characterized by a high concentration of helium-3 to helium-4, and possibly other primitive isotopic features[10]. The authors speculate that this high-nickel-content source contributes to upwelling mantle plumes that originate near the core–mantle boundary, and that it is tapped when the plumes melt near Earth's surface.

Herzberg and colleagues' model is

---

*This article and the paper under discussion[3] were published online on 9 January 2013.

50 Years Ago

In arboriculture the word 'mulch' is usually associated with the compost of wet leaves, straw, etc., laid as a protection to the roots of newly planted or young trees, especially on inherently dry soils … In the autumn number of the *Esso Magazine* (1962), an informative and well-illustrated account of a petroleum mulch evolved by Esso Research is given … Briefly, this type of mulch is in fact an inexpensive emulsion of petroleum resins, used as an agricultural spray for improving the growth and yield of crops of all kinds. Its function is four-fold: to warm the soil, to reduce evaporation, to protect against soil-erosion, and to retard the dispersal by rain-wash and other agents of vital mineral and chemical constituents in the soil growth-zones wherever it is used … Tests on cotton plants in Arizona in 1960 revealed that where this mulch was used, twelve days after planting, 121,900 plants per acre had emerged as against 8,710 per acre on the control area. The final yield in bales per acre was 2.2 with mulch and 1.65 without, an increase of 34 per cent.
From *Nature* 19 January 1963

100 Years Ago

Sir Ray Lankester's weekly contributions to *The Daily Telegraph* represent the high-water mark of popular papers on scientific subjects. The general public has in recent years been infected with a feverish desire for sensation; and as science can offer little to gratify that appetite, thoughtful articles upon its achievements are now relatively much fewer in the periodical Press than they were a generation or two ago … The papers are perfect models of scientific exposition: simple, yet not childish; informative, but not tedious; bright without being flippant; sparkling with human interest and original always.
From *Nature* 16 January 1913

---

provocative and requires further verification. For example, additional experiments are needed to clarify the role of temperature in the partitioning of nickel between olivine and the melt. Also needed are experiments performed under the conditions present at the core–mantle boundary, because this could elucidate how siderophile elements partition between molten mantle and liquid core in these extreme circumstances. Modelling of the abundances of siderophile elements other than nickel in primary melts, such as cobalt, may provide further tests of the deep-mantle-reservoir model. Finally, the composition of the magmas studied should be investigated using a range of radiogenic isotope tracers (for example, lead, neodymium and hafnium), to further analyse the primordial nature of this potential mantle source rock[10]. ■

Michael J. Walter *is in the Department of Earth Sciences, University of Bristol, Bristol BS8 1RJ, UK.*
e-mail: m.j.walter@bristol.ac.uk

1. Hilton, D. R. & Porcelli, D. in *The Mantle and Core, Treatise on Geochemistry* Vol. 2 (ed. Carlson, R. W.) 277–318 (Elsevier, 2005).
2. McDonough, W. F. & Sun, S. *Chem. Geol.* **120,** 223–253 (1995).
3. Herzberg, C. *et al. Nature* **493,** 393–397 (2013).
4. Li, C. & Ripley, E. M. *Chem. Geol.* **275,** 99–104 (2010).
5. Putirka, K., Ryerson, F. J., Perfit, M. & Ridley, W. I. *J. Petrol.* **52,** 279–313 (2011).
6. Sobolev, A. V., Hofmann, A. W., Sobolev, S. V. & Nikogosian, I. K. *Nature* **434,** 590–597 (2005).
7. Sobolev, A. V. *et al. Science* **316,** 412–417 (2007).
8. Labrosse, S., Hernlund, J. W. & Coltice, N. *Nature* **450,** 866–869 (2007).
9. Walter, M. J. & Trønnes, R. G. *Earth Planet. Sci. Lett.* **225,** 253–269 (2004).
10. Jackson, M. G. *et al. Nature* **466,** 853–856 (2010).

STEM CELLS

Surf the waves of reprogramming

Cellular reprogramming to a stem-cell-like state is inefficient and poorly understood, despite its biomedical potential. Detailed molecular analyses of this process are now reported, and should help to overcome these limitations.

IGNACIO SANCHO-MARTINEZ & JUAN CARLOS IZPISUA BELMONTE

Induced pluripotent stem cells can generate all tissues of the adult organism and — unlike embryonic stem cells — their generation does not have ethical implications[1]. They are thus invaluable for research, and have great potential for use in regenerative medicine and drug discovery. The precise molecular mechanisms that underlie the reprogramming of differentiated cells to induced pluripotent stem cells remain largely unknown. Three papers[2–4], published in *Cell* and *Cell Reports*, provide the most comprehensive analysis yet of the molecular dynamics that bring about this process. This work may not only guide the design of improved reprogramming approaches, leading to increased efficiency, but also allow the generation of higher-quality stem cells. Most intriguingly, we think that the demonstration that the cells transit through distinct intermediate states is an elegant indication that reprogramming proceeds in an inverse manner to the process of embryonic development.

In addition to the low efficiency of existing protocols for reprogramming to induced pluripotent stem (iPS) cells, a lack of molecular markers hinders the monitoring of early stages of this process. Consequently, to unravel the molecular basis of reprogramming, researchers have focused on the initial and late phases of the reprogramming process — until now.

In a comprehensive single-cell study, Buganim *et al.*[2] investigated the dynamics of the total cellular transcripts (the transcriptome) underlying reprogramming. The authors analysed reprogramming-competent cells, which they could distinguish from other partially reprogrammed cells. They identified a hierarchical genetic network, which orchestrated the acquisition of pluripotency, and found previously unknown reprogramming factors. Meanwhile, Polo *et al.*[3] and Hansson and colleagues[4] have demonstrated that reprogramming dynamics can be monitored in bulk cell populations. By comparing single-cell data with cell-population results, Polo and co-workers could distinguish cells poised to become iPS cells from reprogramming-resistant populations.

Those authors' findings[3] indicate that, during reprogramming, cells transition through intermediate states that ultimately lead to the appearance of a population expressing Oct-4. (Oct-4, along with three other transcription factors, Sox2, c-Myc and Klf4, mediate reprogramming to iPS cells and are known as the Yamanaka factors[1].) It is with the onset of Oct-4 expression that cells generally become capable of generating iPS cells.

Notably, further increase in the expression of Yamanaka factors could rescue — at least to some extent — cells that were refractory to reprogramming. This suggests that reprogramming-resistant and reprogramming-competent populations express different levels of these factors, and it provides insight into the specific roles that the various Yamanaka factors have during reprogramming-associated cell transitions. For instance, c-Myc was found to be the driving force during the early phase, whereas Oct-4 and Sox2 played a crucial part during the intermediate phases. Klf4 was especially active during the early and late phases, with reduced activity during the intermediate phase[3].

The changes in the transcriptome that underlie cellular transitions occurred in two main 'waves' (Fig. 1). The first wave included differential expression of genes involved in the cell-division cycle, DNA replication and a process called the mesenchymal-to-epithelial transition, which also occurs during normal development. In parallel, some cells showed reductions in the expression of genes associated with cell–cell interaction and cell adhesion, and of markers that are typical of the initial differentiated-cell population.

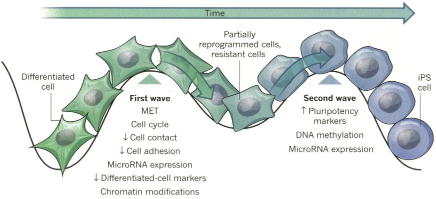


Figure 1 | Waves of molecular events underlie cellular reprogramming[2–4]. Reprogramming of differentiated cells to induced pluripotent stem (iPS) cells involves two main waves of molecular-remodelling events. During the first wave, differentiated cells undergo transcriptome and chromatin changes, which affect the mesenchymal–epithelial transition (MET), among other cellular and molecular phenomena, and which lead to erasure of the differentiated-cell identity. The two waves are separated by an intermediate phase of gradual changes that lead to increased expression of genes involved in early stages of the transition to pluripotency. In this phase, reprogramming-refractory and reprogramming-competent cells coexist, giving rise to more heterogeneous cell populations. During the second wave, cells hierarchically increase the expression of genes that are involved in the establishment and maintenance of pluripotency, and changes to DNA methylation help to lock the reprogrammed state.

An intermediate phase of gradual changes then followed, preceding the second wave. These included a progressive increase in the expression of early — albeit not definitive — markers of pluripotency, including the proteins Fbxo15 and alkaline phosphatase. In the second wave, genes that are known to mediate the establishment and maintenance of pluripotency were increasingly expressed in a hierarchical manner.

Remarkably, Polo *et al.* also detected two temporally distinct cascades of events at the regulatory or epigenetic levels. (Epigenetic changes are chemical changes to DNA, or to nuclear proteins, that affect gene expression.) Regulation of microRNA sequences, which mediate control of gene expression, occurred in two waves that were inversely correlated with the expression of the microRNAs' expected target genes. With regard to epigenetic modifications, histone proteins — which wrap around DNA to form chromatin — were modified during the early phase of reprogramming, whereas changes to DNA-methylation patterns occurred exclusively in the late stages. This sequence of molecular events is reminiscent of cell differentiation, in which transcriptome and chromatin changes occur before changes in DNA methylation, thereby 'locking' the cells in the differentiated state. This finding seems to contradict the previously mentioned cell-transition events, which seemingly occur in the opposite direction to development. Nonetheless, in both cases, DNA-methylation changes occur first and contribute to locking a defined cellular state — whether it is the pluripotent or the differentiated state.

Using the same approach, Hansson *et al.*[4] carried out a dynamic study of protein expression during reprogramming to iPS cells. They isolated different cell populations at different times during this process to analyse the complement of proteins differentially regulated (the proteome), and found that the genetic and epigenetic changes that Polo and co-workers describe are reflected in two highly coordinated waves of regulation of protein expression.

These results[3,4] open up unprecedented possibilities for detailed investigation of the reprogramming process and for ways to optimize it. Nonetheless, more work needs to be done to establish whether reprogramming of human cells follows a similar molecular roadmap to the present studies in mouse cells. Moreover, changes observed at the chromatin level cannot be explained by a direct effect of Yamanaka factors, as these transcription factors are not known to directly mediate chromatin remodelling, and so the causes of such changes should be determined. Furthermore, it is unclear whether Yamanaka factors directly regulate the expression of all differentially regulated genes, particularly during the early phase, when massive changes rapidly occur on introduction of the genes that encode these factors into the cell.

Another intriguing question relates to cell-lineage conversion processes that are triggered by transient expression of the Yamanaka factors, without the need for initial reprogramming to a pluripotent state[5]. The observation that these regulators of development are themselves dynamically regulated during reprogramming to iPS cells suggests that reprogramming might generate intermediate states with different differentiation potentials, as is observed during embryonic development. Thus, further study of the molecular dynamics underlying the different reprogramming processes might provide insight into the fundamental mechanisms that define cellular 'plasticity' and that might facilitate the generation of cell populations for clinical applications.

Together, these observations constitute the most detailed description of reprogramming dynamics at the transcriptome, epigenome and proteome levels reported so far. They also demonstrate that sorted cell populations are suitable for molecular studies of this process, and that their analysis perfectly complements single-cell studies.

These three studies not only identify further reprogramming factors, but also could serve as the basis for the development of highly efficient reprogramming strategies. This, in turn, might allow large numbers of iPS cells to be rapidly generated, which could contribute to a greater iPS cell safety profile by enabling more detailed investigation of the cells.

Finally, the observation that reprogramming intermediates might be correlated with different developmental states should permit more precise studies of cellular differentiation. And the knowledge gained by an in-depth epigenetic analysis of intermediate cell populations may aid the translation of reprogramming strategies from the lab to the clinic. ■

Ignacio Sancho-Martinez and **Juan Carlos Izpisua Belmonte** *are at the Salk Institute for Biological Studies, La Jolla, California 92037, USA, and at the Center of Regenerative Medicine, in Barcelona, Spain.*
e-mail: belmonte@salk.edu

1. Takahashi, K. & Yamanaka, S. *Cell* **126,** 663–676 (2006).
2. Buganim, Y. *et al. Cell* **150,** 1209–1222 (2012).
3. Polo, J. M. *et al. Cell* **151,** 1617–1632 (2012).
4. Hansson, J. *et al. Cell Rep.* **2,** 1579–1592 (2012).
5. Sancho-Martinez, I. *et al. Nature Cell Biol.* **14,** 892–899 (2012).

EVOLUTIONARY GENETICS

Genes for home-building

Dissecting the genetic basis of a species' physical features is regular research fodder. So why not do the same for the inanimate objects that some animals produce, such as beehives or beaver dams? On page 402 of this issue, Weber and colleagues show that the sophisticated burrows of oldfield mice can be understood using straightforward genetics, shedding light on how this classic 'extended phenotype' evolved (J. N. Weber *et al. Nature* **493,** 402–405; 2013).

Oldfield mice (*Peromyscus polionotus*) and their relatives dig homes that, among other things, are used to avoid predators, and different species produce distinct shapes and sizes of burrow. These easily measurable physical manifestations of complex behaviours are a handy way of asking questions about the genetics that underpins those behaviours.

The burrows of oldfield mice (pictured) are of a consistent size, and have an entrance tunnel leading to the nest, together with an exit tunnel that lies just below the surface and allows escape if a predator comes to the front door. By contrast, the related deer mouse (*Peromyscus maniculatus*) has smaller burrows that have only an entrance tunnel. It is thought that this simpler burrowing style is the ancestral state, and that oldfield mice evolved the more complex strategy when they moved into the open habitats that they occupy.

Weber *et al.* find that mice kept in laboratory conditions dig burrows just as they do in the wild and that, when the two species are crossed, the offspring dig oldfield-style tunnels, indicating that this evolutionarily newer behaviour is genetically dominant.

VERA DOMINGUES

Backcrosses between the hybrids and the parental species allowed the authors to identify the genetic regions controlling burrow properties. They found that entrance-tunnel length is a complex trait governed by at least three chromosomal locations, or loci, whereas the presence of an escape tunnel is governed by a single, separate locus.

These results suggest that complex behaviours can arise from a combination of simpler genetic modules that have evolved separately. **Patrick Goymer**

NEUROSCIENCE

Memory and the single molecule

Sustained activity of the brain-specific enzyme PKM-ζ is thought to underlie the maintenance of long-term memories. Studies in PKM-ζ-deficient mice, however, cast the importance of this protein into question. SEE LETTERS P.416 & P.420

PAUL W. FRANKLAND & SHEENA A. JOSSELYN

Memory traces are thought to be formed by the strengthening of synaptic connections between particular collections of neurons. A major goal of neuroscience research is to identify the molecular machinery that sustains these strengthened connections once they are formed. Many molecules that were once thought to have starring roles in memory have since been relegated to supporting, albeit still important, roles. Two papers in this issue (Volk *et al.*[1] and Lee *et al.*[2])* challenge the star status of yet another memory molecule — protein kinase M-ζ.

The human genome encodes more than 500 protein kinase enzymes, and many of these are known to mediate memory formation. Protein kinase M-ζ (PKM-ζ) is special in that it is proposed to maintain the enhanced synaptic strength associated with memory formation — unlike most other kinases, which are involved in synaptic strengthening per se. Moreover, in contrast to other kinases, PKM-ζ is always active, a perfect feature for a molecule that retains memory. It has therefore been considered truly deserving of 'memory molecule' billing.

One of the first hints of the involvement of PKM-ζ in sustaining synaptic strength came from a study[3] in which the function of the enzyme was disrupted after induction of long-term potentiation (LTP, a form of synaptic strengthening that is often used as a cellular proxy for memory formation) in the brain's hippocampus region. This investigation found that both an inhibitory form of PKM-ζ and a pharmacological inhibitor of this protein called ZIP reversed previously established LTP. (ZIP stands for zeta inhibitory peptide and is a 13-amino-acid sequence thought to mimic the natural substrate that turns PKM-ζ off.)

This finding received considerable attention, but the true red-carpet moment for PKM-ζ came when another paper[4] showed that intrahippocampal microinjections of ZIP — but not of a scrambled ZIP peptide that did not affect PKM-ζ and so was used as a control — reversed not only established LTP but also an established memory in rats. That disrupting the function of a molecule could erase a memory created a great buzz and sparked numerous sequels. Indeed, microinjection of ZIP into many different brain regions was subsequently shown to disrupt established spatial, fear, appetitive, habit and sensorimotor memories in rodents and sensitization memory in the sea slug *Aplysia*[5].

Using different strategies to genetically delete PKM-ζ, Volk *et al.* and Lee *et al.* now challenge the conclusions of these previous studies. They question both the specificity of ZIP and the importance of PKM-ζ in maintaining LTP and memory. Volk and colleagues created a line of mice in which *Pkm-ζ* and a closely related gene, *Pkc-ζ*, were deleted throughout the animal in embryonic development. Despite having no PKM-ζ, these mice seemed normal, and their hippocampal LTP was indistinguishable both in magnitude and duration from LTP in normal mice.

Because absence of a gene since early development may result in compensation for its function by other genes, the authors also created a line of mice in which PKM-ζ could be deleted specifically in excitatory forebrain neurons in adult animals. These mice similarly showed normal LTP. The authors went on to perform a clever experiment in which they examined the effects of ZIP on maintenance of LTP in mice that entirely lacked PKM-ζ. Remarkably, ZIP (and scrambled ZIP), at doses used in previous experiments, disrupted established LTP not only in normal mice, but also in PKM-ζ-deficient mice. This result indicates that ZIP disrupts established LTP by PKM-ζ-independent mechanisms and brings into question the role of PKM-ζ in the maintenance of LTP.

---

*This article and the papers under discussion[1,2] were published online on 2 January 2013.

Lee *et al.* also created mice that lacked PKM-ζ from early in development and, using an exhaustive battery of behavioural tests, probed their memory. The mice could form persistent memories in the various tests — an observation that is independently reported by Volk and co-workers. Lee *et al.* also examined the effects of ZIP on existing memories. For this, they microinjected ZIP into the brain of PKM-ζ knockout mice after formation of a place memory. As a striking counterpart to the LTP results of Volk *et al.*, ZIP erased memory in mice lacking PKM-ζ. Although the specificity of ZIP (and of scrambled ZIP) has been debated[6,7] and the jury is still out on the precise molecular targets of these peptides, these data convincingly show that ZIP disrupts both established LTP and memory independently of PKM-ζ.

These results, however, do not entirely exclude the possibility that PKM-ζ is a key player in memory maintenance. They do not account for two previous studies in which genetic tools were used to manipulate PKM-ζ function following learning. One of these studies[8] found that acute expression of a mouse PKM-ζ transgene in the fruitfly *Drosophila* 30 minutes after training enhanced odour memory measured one or four days after training. The second study[9] showed that microinjecting a lentivirus expressing PKM-ζ into the brain of rats six days after training enhanced a taste-aversion memory. It also found that disrupting PKM-ζ function by inducing expression of an inhibitory form of the enzyme days after training was sufficient to erase an established taste-aversion memory.

How might these findings[1,2,8,9] be reconciled? Hundreds of molecules are probably involved in the formation and maintenance of memory, and within this 'mnemome' there may be a fair degree of redundancy and degeneracy — whereby related molecules compensate for each others' deficiency[10]. Such mechanisms could make up for long-term deficiency of a single molecule, allowing formation of stable memories. Acute, post-training manipulations may reduce the likelihood of this type of compensation but might not eliminate it.

So asking whether PKM-ζ could really be 'the' memory molecule is probably not the right question, as it fails to capture the complexity of the interactions between the molecules within the mnemome. Given evolution's penchant for redundancy, it seems unlikely that any single molecule will play this part solo. Rather, the question should be one of degree — is a molecule a lead player (and therefore irreplaceable) or does it have more of a supporting role? The casting debate on PKM-ζ will, no doubt, continue. ■

Paul W. Frankland *and* **Sheena A. Josselyn** *are in the Program in Neurosciences and Mental Health, Hospital for Sick Children, Toronto, Ontario M5G 1X8, Canada. They are also in the Departments of Psychology and of Physiology and at the Institute of Medical Science, University of Toronto.*
e-mail: paul.frankland@sickkids.ca

1. Volk, L. J., Bachman, J. L., Johnson, R., Yu, Y. & Huganir, R. L. *Nature* **493,** 420–423 (2013).
2. Lee, A. M. *et al. Nature* **493,** 416–419 (2013).
3. Ling, D. S. F. *et al. Nature Neurosci.* **5,** 295–296 (2002).
4. Pastalkova, E. *et al. Science* **313,** 1141–1144 (2006).
5. Sacktor, T. C. *Nature Rev. Neurosci.* **12,** 9–15 (2011).
6. Wu-Zhang, A. X., Schramm, C. L., Nabavi, S., Malinow, R. & Newton, A. C. *J. Biol. Chem.* **287,** 12879–12885 (2012).
7. Yao, Y. *et al. Neuropharmacology* **64,** 37–44 (2013).
8. Drier, E. A. *et al. Nature Neurosci.* **5,** 316–324 (2002).
9. Shema, R. *et al. Science* **331,** 1207–1210 (2011).
10. Sanes, J. R. & Lichtman, J. W. *Nature Neurosci.* **2,** 597–604 (1999).

ANALYTICAL CHEMISTRY

Defects visualized in porous solids

The use of confocal fluorescence microscopy to image defects in crystalline, porous solids known as metal–organic frameworks enables the relationship between the number of defects and the materials' properties to be determined.

CHRISTIAN SERRE

Porous crystalline solids are of great interest for many chemical applications, including separations, catalysis and ion exchange[1]. Such solids often contain defects that can markedly impair their performance[2], but visualizing the defects in single particles of material can be difficult. Writing in *Angewandte Chemie*, Ameloot *et al.*[3] report that an optical imaging technique called confocal fluorescence microscopy can be used for the three-dimensional visualization of defects in single crystals of metal–organic frameworks — a versatile class of porous material.

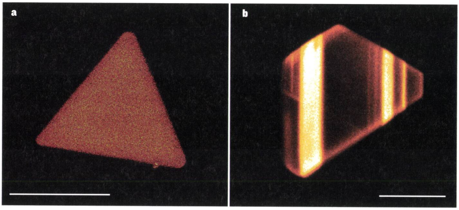
Metal–organic frameworks (MOFs) are crystalline solids that are formed when metal cations strongly associate with organic linker molecules[4,5] to generate a highly porous network. By varying the building blocks and the conditions used to synthesize these solids, a wide range of materials with different structural and chemical features can be made, often leading to high porosities and tunable pore sizes (ranging from a few ångströms to several nanometres). Many MOFs can be modified by attaching chemical groups to the linkers, and they sometimes possess a flexible pore system that can adapt to the presence of guest molecules by changing shape.

The high porosity and tunable composition of MOFs represent key advantages over the material's wholly organic or inorganic porous counterparts. But they are often more fragile than their counterparts when exposed to heat or chemicals, because the bonds between the ions and the linkers tend to break fairly easily[6]. They are therefore prone to developing defects, which means that the performance of MOFs for a given application can be more variable than that of other classes of porous material.

The relationship between defect formation and MOF performance would be more easily established if the defects in these materials could be visualized. Such an analysis can be performed using several techniques, such as X-ray diffraction or scanning probe microscopy. But these methods either cannot provide insight at every stage of defect formation or suffer from limitations, such as revealing defects only at the outer surfaces of particles.

Confocal fluorescence microscopy provides micrographs of sufficient resolution and contrast to image defects, and enables three-dimensional images to be constructed from a series of two-dimensional images of cross-sections through an object[7]. Ameloot *et al.* have used the technique to observe defect formation in three dimensions in metal polycarboxylate MOFs (in which the linkers bind to metals using carboxylate groups, CO_2^-). Such MOFs often have defects in which broken bonds between cations and carboxylates generate free carboxylic acid groups (COOH) within pores.

The authors' approach relies on the adsorption and diffusion of a small molecule, furfuryl alcohol, within the MOFs. Furfuryl alcohol is



REF. 3

Figure 1 | Defect accumulation in metal–organic frameworks. Ameloot *et al.*[3] have used confocal fluorescence microscopy to image defects in single crystals of porous materials known as metal–organic frameworks (MOFs). **a,** In this image of the triangular surface of a MOF crystal that was prepared in a short synthesis time (15 hours), the lack of any bright features indicates an absence of defects. **b,** In this cross-section of a crystal made of the same material as **a**, but which was prepared using a longer synthesis time (63 h), defects are seen as bright lines. Scale bars, 10 micrometres.

easily adsorbed within pores, but cannot break the metal–carboxylate bonds of the MOF structure. Crucially, it also undergoes polymerization when it encounters acids, a reaction that generates fluorescence. Ameloot *et al.* therefore used confocal fluorescence microscopy to follow the polymerization of furfuryl alcohol at acidic sites (defects) within the pores of the MOF.

The researchers tested their method by analysing a well-studied MOF known as HKUST-1 (ref. 8), choosing two samples that apparently had the same crystallinity and porosity, but which were made using different synthesis times — one sample was made in a reaction that ran for a short time (15 hours), whereas the other was allowed much longer (63 h). The authors' confocal analysis revealed almost no defects in the sample made in a short time, but a high defect content in the other sample (Fig. 1). Notably, the technique not only detected defects, but also identified the main crystallographic direction in which the defects formed.

Ameloot *et al.* went on to illustrate how defect content affects the properties of their porous MOFs. They examined the catalytic performance of the solids in a reaction that depends on the number of carboxylic acid groups in the pores, evaluating samples that had been made using different synthesis times and which therefore had different defect contents. Using their microscopy technique to visualize defects in the samples, they demonstrated that the catalytic activity of the MOF was directly related to the number of defects. The same was true for two other polycarboxylate MOFs, known as MOF-5 and MIL-53(Ga).

If small molecules can be found that produce fluorescent signals in the presence of other chemical groups in MOF networks, confocal fluorescence microscopy might be useful for assessing the defect content of hundreds of other porous MOFs. One can also imagine using the technique to compare the quality of MOFs obtained by different synthetic routes.

Another application would be to evaluate MOF stability — which is associated with defect formation — when the solids are exposed to conditions that typically provoke their degradation, such as heat, water or complex biological fluids. The information obtained might be crucial for the development of applications including gas-storage systems and chemical separations. ■

Christian Serre *is at the Institut Lavoisier Versailles, University of Versailles St-Quentin-en-Yvelines, 78035 Versailles, France.*
e-mail: serre@chimie.uvsq.fr

1. Schüth, F., Sing, K. S. W. & Weitkamp, J. (eds) *Handbook of Porous Solids* (Wiley-VCH, 2002).
2. Feldhoff, A. *et al. ChemPhysChem* **10,** 2429–2433 (2009).
3. Ameloot, R. *et al. Angew. Chem. Int. Edn* **52,** 401–405 (2013).
4. Kitagawa, S. & Matsuda, R. *Coord. Chem. Rev.* **251,** 2490–2509 (2007).
5. Férey, G. *Chem. Soc. Rev.* **37,** 191–214 (2008).
6. Low, J. J. *et al. J. Am. Chem. Soc.* **131,** 15834–15842 (2009).
7. Müller, M. A. *Introduction to Confocal Fluorescence Microscopy* 2nd edn (SPIE Press, 2006).
8. Chui, S. S.-Y., Lo, S. M.-F., Charmant, J. P. H., Orpen, A. G. & Williams, I. D. *Science* **283,** 1148–1150 (1999).

COMPUTATIONAL MATERIALS SCIENCE

Trustworthy predictions

A method has been developed to compute the precise quantum-mechanical properties of certain insulators. This approach avoids the uncertainties that are intrinsic to predictions made using existing approaches. SEE ARTICLE P.365

PAUL R. C. KENT

One of the great promises of computational materials science, physics and chemistry is the ability to predict the properties of materials accurately and reliably. Knowledge of a material's elemental composition and some hints of its structure should be sufficient to predict and explain the material's intrinsic strength, for example, or its suitability for catalysis or energy storage, and whether it displays any unusual physical phenomena. In other words, researchers in these fields hope to combine the rational design of materials with the means to accurately determine how they function. On page 365 of this issue, Booth *et al.*[1] report a computational method that brings this goal a step closer*.

Fundamentally, all the above properties are determined by the way electrons in a material are bound to atoms. Electron behaviour is quantum mechanical and described by the Schrödinger equation, in which every electron interacts with every other electron. This leads to combinatorial complexity and the 'many-body problem': solving the equation accurately and reliably is a grand challenge, even using the largest supercomputers. In practice, researchers must make considerable approximations to study realistic materials computationally, adding uncertainty to their predictions. It can therefore be difficult to trust unexpected results, or to determine whether the right answers are being obtained for the correct reasons.

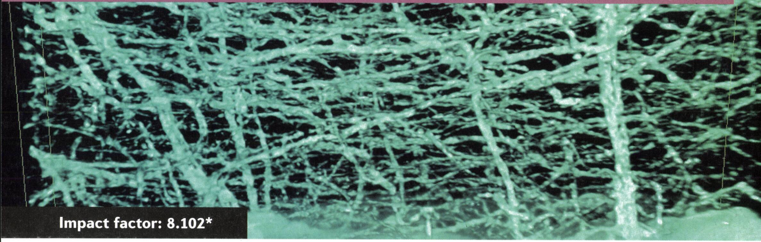
Booth and colleagues' computational method solves the Schrödinger equation without the need for any approximations. The authors show that accurate energies for electrons and atoms can be obtained for at least a small class of insulating materials. They go on to demonstrate that a computational method that is widely accepted as providing the benchmark for theoretical predictions of molecular properties can accurately predict the properties of some materials. These advances will help to establish a firm foundation for quantum-mechanics-based simulations of material properties, in which the most

---

*This article and the paper under discussion[1] were published online on 19 December 2012.

nature REVIEWS GASTROENTEROLOGY & HEPATOLOGY

Collection on the Enteric Nervous System



Impact factor: 8.102*

A renaissance in the field of neurogastroenterology has enabled advances in our understanding of the function of the enteric nervous system in both gastroenterological and neurological disorders. This collection showcases recent progress in the field and highlights the promise of new treatments for functional gastrointestinal diseases.

All content for this Collection has been chosen by the *Nature Reviews Gastroenterology & Hepatology* editors and the collection has been made freely available for 6 months thanks to support from Ferring.

FREE online for a limited time: www.nature.com/nrgastro/collection/enteric

Scan here with your smart phone to access the focus.
Download the free app for your phone at http://gettag.mobi

Image credit: Courtesy of S.-C. Tang
*2011 *Journal Citation Reports*® (Thomson Reuters, 2012).
Nature Reviews Gastroenterology & Hepatology was previously published as *Nature Clinical Practice Gastroenterology & Hepatology*

Produced with support from: FERRING PHARMACEUTICALS

nature publishing group npg

important approximations can be assessed quantitatively.

The dominant approach for quantum-mechanics-based materials simulation is density functional theory[2] (DFT), which maps the many-body problem to an equivalent problem in which the electrons do not interact directly with each other. In principle, this leads to a simpler, but still exact, solution of the Schrödinger equation. However, the exact mapping (the 'functional') is not known, so approximations are required. Fortunately for researchers, the available approximations are good enough for them to model the properties of many materials, yielding predictions within a few per cent of values obtained experimentally. There are some major problems with DFT, however, including its inability to correctly reproduce the energy curve that describes the binding of a single electron shared between two protons[3]; the exact solution of this problem is easy to obtain, because with only one electron, there is no many-body problem to solve.

Despite the thousands of papers published each year in which DFT is used as a modelling tool, a fundamental paradox remains: because the methods of DFT are not exact, any predictions must be carefully rationalized against experimental results and against the properties of similar materials. The methods are therefore at their weakest when little is known about the material being modelled. Alternative approaches are needed that are exact and that systematically converge on the experimental result. Such an approach would inevitably be more computationally expensive than DFT, but the costs could be justified by the savings made from having to do less experimental validation, and by the potential scientific value of the findings.

Booth and co-workers' method — full configuration interaction quantum Monte Carlo (FCIQMC) — uses a statistical approach to solve the Schrödinger equation. Other statistical methods have previously been used to solve the Schrödinger equation for materials, but these required approximations or have not been demonstrated to work well for a wide range of materials. The version of FCIQMC used to predict molecular properties is only 3 years old[4], but it has been rapidly developed and has in the past year been used to calculate the optical properties[5] of molecules, for example.

The method tames the exponential complexity of the many-body problem to such an extent that useful results can be obtained using current computational resources. Booth *et al.* describe extensions to the method that enable it to calculate the properties of solids that have periodic arrangements of atoms. The authors demonstrate that sufficiently converged properties can be achieved — albeit with some difficulty — for some insulating materials. Until now, no exact methods have been available to predict the properties of even the simplest of solids, although some approaches have come close[6,7].

Using FCIQMC can require tens of thousands of computation hours, whereas a comparable DFT calculation might take one hour. However, if the computed properties are converged, the FCIQMC result will be exact within statistically defined error bars, whereas the error for the DFT result will be unknown. One use of FCIQMC may therefore be to help to improve approximate methods for calculating material's properties.

In this spirit, the authors tested whether the gold-standard approach used by computational chemists to predict the properties of molecules — a method known as CCSD(T) — could be extended to predict the properties of solids, by comparing its results with those from FCIQMC. Promisingly, the researchers found that the errors of the CCSD(T) results were much smaller than those of many modern DFT functionals. Although the number of calculations required for CCSD(T) scales with the seventh power of the number of atoms, a wealth of computational-chemistry techniques and efficient implementations may reduce the practical cost.

> ***"Booth and colleagues' computational method solves the Schrödinger equation without the need for any approximations."***

One major challenge will be to extend the applicability of the FCIQMC method to larger and more complex systems than those studied by Booth and colleagues. The oxides of the first row of transition metals in the periodic table would be good targets, for example, because they display structural, magnetic and electronic phenomena that are challenging to reproduce in existing computational models. Even if full convergence of the predicted properties cannot be achieved, information about the quantum-mechanical wavefunction of these compounds may be highly informative.

For now, though, Booth and co-workers' method is restricted to simulating very simple materials at a temperature of absolute zero. Improvements to the algorithm that enhance performance can be expected — even a doubling of the system size that can be modelled would be useful. One can also envisage the development of an 'embedding' method that incorporates FCIQMC results into those of a cheaper approach to increase the achievable length scales. This would give accurate results if the physical effects that determine properties of interest are strongly localized. But surfaces and defects in solids will be difficult to model using the authors' approach because they require large numbers of atoms to be considered.

Promisingly, metals might be studied using modest extensions to Booth and colleagues' method, or by using the current method in combination with brute-force application of computer power. This would enable the entire phase diagram (a graph plotting the physical state of a material at different pressures and compositions) of simple materials to be predicted with confidence. For example, DFT predictions of previously unknown phases of elemental solids have been made in the past few years[8–10]. In future, FCIQMC might be able to determine exactly the pressures at which these phases occur, removing lingering suspicions around the accuracy of the DFT results.

For Booth and co-workers' method to have the greatest impact, researchers will need to learn to interpret the results of FCIQMC calculations and use them to improve cheaper, everyday methods for modelling materials properties. The common language of FCIQMC and computational chemistry may enable the development of specialized methods for materials, with an accuracy and reliability far exceeding today's DFT approaches. By helping to establish a hierarchy of trusted and validated computational methods for materials, the new technique should help to deliver one of the great promises of computational materials science. ■

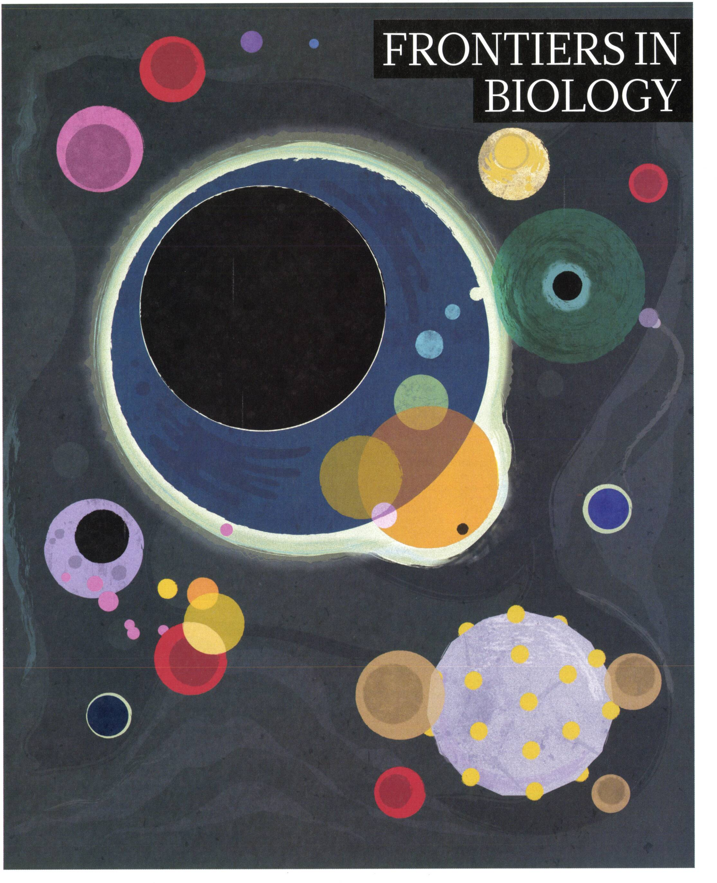
Paul R. C. Kent *is at the Center for Nanophase Materials Sciences and the Computer Science and Mathematics Division, Oak Ridge National Laboratory, Oak Ridge, Tennessee 37831-6494, USA.*
e-mail: kentpr@ornl.gov

1. Booth, G. H., Grüneis, A., Kresse, G. & Alavi, A. *Nature* **493,** 365–370 (2013).
2. Martin, R. M. *Electronic Structure: Basic Theory and Practical Methods* (Cambridge Univ. Press, 2004).
3. Cohen, A. J., Mori-Sánchez, P. & Yang, W. *Science* **321,** 792–794 (2008).
4. Booth, G. H., Thom, A. J. W. & Alavi, A. *J. Chem. Phys.* **131,** 054106 (2009).
5. Booth, G. H. & Chan, G. K.-L. *J. Chem. Phys.* **137,** 191102 (2012).
6. López-Ríos, P., Ma, A., Drummond, N. D., Towler, M. D. & Needs, R. J. *Phys. Rev. E* **74,** 066701 (2006).
7. Purwanto, W., Krakauer, H. & Zhang, S. *Phys. Rev. B* **80,** 214116 (2009).
8. Pickard, C. J. & Needs, R. J. *Nature Mater.* **7,** 775–779 (2008).
9. Ma, Y. *et al. Nature* **458,** 182–185 (2009).
10. Pickard, C. J. & Needs, R. J. *Nature Mater.* **9,** 624–627 (2010).

CORRECTION

In the News & Views article 'Astronomy: Andromeda's extended disk of dwarfs' by R. Brent Tully (*Nature* **493,** 31–32; 2013), 13 dwarf companion galaxies of the Andromeda galaxy are described as lying "at distances from Messier 31 of between 35 and 400 kiloparsecs (114–1,305 light years)". The distance values in units of light years should have been 114,000 to 1.3 million.

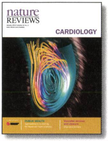
nature INSIGHT



FRONTIERS IN BIOLOGY

CARDIOLOGY | ENDOCRINOLOGY | GASTROENTEROLOGY & HEPATOLOGY | NEPHROLOGY | NEUROLOGY | CLINICAL ONCOLOGY | RHEUMATOLOGY | UROLOGY

This month in *Nature Reviews*



Myofibroblast-mediated mechanisms of pathological remodelling of the heart
K. T. Weber, Y. Sun, S. K. Bhattacharya, R. A. Ahokas & I. C. Gerling

In this Review, Weber and colleagues discuss mechanisms of cardiac fibrosis and consequent adverse remodelling of the heart. They also highlight potential strategies for cardioprotection.

www.nature.com/nrcardio

nature REVIEWS
ENDOCRINOLOGY

Global obesity: trends, risk factors and policy implications
V. S. Malik, W. C. Willett & F. B. Hu

In this review, the influence of globalization on global trends in obesity prevalence and population-based interventions to combat the obesity epidemic are discussed and broad policy recommendations for obesity prevention are made.

www.nature.com/nrendo

nature REVIEWS
GASTROENTEROLOGY & HEPATOLOGY

New treatments for IBS
M. Halland & N. J. Talley

IBS is a common disorder that can have an enormous negative effect on patients and society as a whole; however, many patients do not achieve adequate symptom relief with the currently available therapies. In this Review, Magnus Halland and Nicholas Talley discuss the latest treatments for IBS, including novel nonpharmacological and pharmacological approaches.

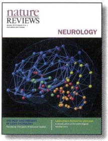
www.nature.com/nrgastro

nature REVIEWS
NEPHROLOGY

Assessing risk in chronic kidney disease: a methodological review
M. E. Grams & J. Coresh

This Review outlines the fundamentals of risk prediction, including considerations pertinent to chronic kidney disease, common methodological shortcomings in risk prediction studies, and metrics often used to assess the performance of risk prediction models.

www.nature.com/nrneph



100 years of Lewy pathology
M. Goedert, M. G. Spillantini, K. Del Tredici & H. Braak

In 1912, Fritz Heinrich Lewy identified the intracellular inclusions that are characteristic of Parkinson disease. Goedert and colleagues review Lewy's life and work, including the discovery of the inclusion bodies that bear his name.

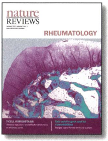
www.nature.com/nrneurol

nature REVIEWS
CLINICAL ONCOLOGY

Clinical opportunities and challenges in targeting tumour dormancy
J. A. Hensel, T. W. Flaig & D. Theodorescu

Tumour dormancy is when cancer sleeps undetected for periods that can last up to decades. This Review describes the mechanisms of dormancy and uses genitourinary cancers as models to demonstrate how dormancy principles could be exploited clinically.

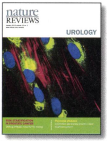
www.nature.com/nrclinonc



siRNA-based therapeutic approaches for rheumatic diseases
F. Apparailly & C. Jorgensen

Small interfering RNAs (siRNAs) are a useful experimental tool to silence gene activity. The challenge, however, remains to optimize delivery and imbue 'drug-like' properties to siRNAs for therapeutic use, as discussed here.

www.nature.com/nrrheum



Paediatric genitourinary cancers and late effects of treatment
K. T. Sadak, M. L. Ritchey & J. S. Dome

Long-term survival of patients with the most common childhood genitourinary cancers are excellent, highlighting the need to minimize the late complications of treatment. Here, the authors discuss the late effects of treating these cancers and consider the need for surveillance of childhood cancer survivors.

www.nature.com/nrurol

nature
REVIEWS

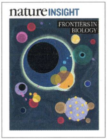
www.nature.com/reviews

nature publishing group npg

natureINSIGHT

FRONTIERS IN BIOLOGY

17 January 2013 / Vol 493 / Issue No 7432



Cover illustration
Nik Spencer

Editor, *Nature*
Philip Campbell
Publishing
Nick Campbell
Claudia Deasy
Production Editor
Jenny Rooke
Art Editor
Nik Spencer
Sponsorship
Reya Silao
Production
Emilia Orviss
Marketing
Elena Woodstock
Hannah Phipps
Editorial Assistant
Anastasia Panoutsou

The *Nature* Insight 'Frontiers in biology' aims to cover timely and important developments in the broader field of biology, ranging from the subcellular to the organismal level, and including molecular mechanisms and biomedicine.

In the first review, Yoshiki Sasai discusses the principles and mechanisms by which organs and tissues form through self-organization from a homogeneous population of cells. He proposes future directions that integrate emerging technologies and mathematical modelling to allow quantitative analysis and a better understanding of self-organization and morphogenesis.

Autism spectrum disorder (ASD) is increasingly considered to be a disorder of synaptic function in neurons. Neural circuits are shaped by experience, and neural activity induces changes in synaptic genes and proteins. Many of the genes associated with ASD are components of activity-dependent neural signalling pathways. Daniel Ebert and Michael Greenberg explore the hypothesis that disruption of this signalling, and hence neural function and plasticity, may be a cause of ASD.

Simon Johnson, Peter Rabinovitch and Matt Kaeberlein then provide an account of research into mTOR's role in ageing and ageing-related disease. mTOR inhibition has been shown to extend lifespan in model organisms and to offer protection against various age-related pathologies. In the review, the authors explore if, when and how drugs that target mTOR could be used to slow ageing.

Highlighting the crucial role of cellular metabolism in regulating immune-cell function and, by extension, inflammatory diseases and autoimmunity, Luke O'Neill and Grahame Hardie discuss the signals and pathways by which immune cells shift their main energy mechanism during inflammation from mitochondrial oxidative metabolism to rapid glucose uptake and glycolysis, and speculate on how dietary restriction or drug-induced pseudo-starvation might help to promote an anti-inflammatory phenotype.

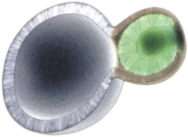
Finally, by describing how mouse models have provided insight into the possible endogenous source of DNA damage that results in the characteristic phenotypes of individuals with Fanconi anaemia, Molly Kottemann and Agata Smogorzewska discuss the role of Fanconi anaemia proteins in genomic stability, cancer susceptibility and stem-cell development.

Francesca Cesari, I-han Chou, Angela K. Eggleston, Marie-Thérèse Heemels, Barbara Marte & Ursula Weiss
Senior Editors

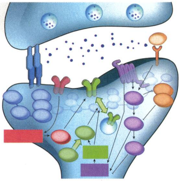
CONTENTS

REVIEWS

318 Cytosystems dynamics in self-organization of tissue architecture
Yoshiki Sasai



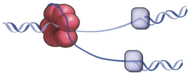
327 Activity-dependent neuronal signalling and autism spectrum disorder
Daniel H. Ebert & Michael E. Greenberg



338 mTOR is a key modulator of ageing and age-related disease
Simon C. Johnson, Peter S. Rabinovitch & Matt Kaeberlein

346 Metabolism of inflammation limited by AMPK and pseudo-starvation
Luke A. J. O'Neill & D. Grahame Hardie

356 Fanconi anaemia and the repair of Watson and Crick DNA crosslinks
Molly C. Kottemann & Agata Smogorzewska



doi:10.1038/nature11859

Cytosystems dynamics in self-organization of tissue architecture

Yoshiki Sasai[1]

Our knowledge of the principles by which organ architecture develops through complex collective cell behaviours is still limited. Recent work has shown that the shape of such complex tissues as the optic cup forms by self-organization *in vitro* from a homogeneous population of stem cells. Multicellular self-organization involves three basic processes that are crucial for the emergence of latent intrinsic order. Based on lessons from recent studies, cytosystems dynamics is proposed as a strategy for understanding collective multicellular behaviours, incorporating four-dimensional measurement, theoretical modelling and experimental reconstitution.

Nature is full of ordered patterns and structures, from the microscopic to macroscopic and even to the cosmic level. Many patterns form from pre-existing ones (prepatterns), such as in DNA replication and cellular proliferation. Patterns with additional complexity can also be generated from a synthesis of multiple prepatterns; for example, mesoderm induction occurs through conjugation of ectodermal and endodermal tissues[1].

Complex patterns and structures can also arise in the absence of obvious prepatterns. Self-organization is the spontaneous formation of ordered patterns and structures from a population of elements (or individuals) that have no or minimal patterns[2–6]. Self-organization is seen in a wide variety of non-living matter on Earth, including molecular assembly in crystal formation and in nanotechnology[7], the Belousov–Zhabotinsky reaction[8], Rayleigh–Bénard convection[9] and even geomorphology of sand ripples[10]. This form of pattern formation is often associated with the concept of emergence, or the spontaneous appearance of an ordered property of the whole that cannot be explained by the sum of the complexity of its elements[2,6]. This has been a matter of intrigue and interest since the time of the Greek philosopher Aristotle, who wrote in his famous work *Metaphysics* "in the case of all things ... in which the totality is not, as it were, a mere heap, but the whole is something besides the parts, there is a cause ...".

Self-organization in life sciences has been studied at multiple scales, from viral particle assembly to the collective behaviours of bees and birds[2,11]. In living and non-living matter, self-organization often has common features[2], such as nonlinear systems, multilayered complex interactions, noise-driven pattern generation and symmetry breaking. In this Review, I focus on structural self-organization at the multicellular level in tissue architecture, and discuss recent studies, new questions and future strategies.

Special matter for tissue self-organization

Dynamic multicellular self-organization seems to have a few aspects that are not commonly seen in physics and chemistry, in which self-organization phenomena have been studied extensively. One of these features is that the 'elements', or cells, radically change along the spatiotemporal axes in a context-dependent manner, for instance, during cellular differentiation. By contrast, snowflakes self-form from molecules of water that remain basically unchanged throughout snowflake formation (Fig. 1a). In addition, biological systems tend to be strongly tied to history or memory, whereby the behaviour of the system is influenced not only by current conditions but also by preceding events. This characteristic history dependency of biological systems — often referred to as stigmergy — is also well known for eusocial behaviours in insects (for example, termites), for which a contextual, environmental and interdependent coordination between individuals that is indirectly affected by their past actions causes a spontaneous emergence of self-organization (for example, the formation of complex termites' nests without centralized instructions)[2,12,13]. Thus, in biological self-organization, interaction rules of elements are not constant but generally evolve in time and space with a history-dependent nature.

In this Review, the term self-organization refers to situations in which an internal system (self) is composed of definable elements that lack prepatterns and can be conceptually distinguished from an external system. Self-organization and extrinsic controls usually co-exist *in vivo*; therefore, I mainly focus on *in vitro* phenomena in mechanistic arguments. In self-organization, the external system may support the operation of the internal system by providing permissive conditions (such as supporting cell growth evenly), but should not give 'biased' information to the cells for patterning. Importantly, a process is considered to be self-organization only when the order of complexity of an autonomously emerging structure substantially exceeds the sum of the complexity in individual cells.

Three modes of tissue self-organization

For simplicity and clarity, the basic processes of tissue self-organization can be classified into the following three categories: self-assembly, self-patterning and self morphogenesis (Fig. 1b). It should be noted, however, that these three processes are not necessarily independent events during tissue development.

Self-assembly

Self-assembly is the time-evolving control of relative cell positions by local rules. Strictly, self-assembly refers to the spontaneous formation of a patterned structure by selective aggregation of cells or by rearrangement of the relative positions of cells within the structure (self-sorting). A complex, highly ordered structure can emerge by self-assembly of just one or a few types of elements[6,14], for example the self-formation of snowflakes. Well-known biological examples at the molecular level are viral shell formation[15] and microtubule assembly[16]. At the tissue level, classic examples are the dissociation–reassembly studies of sea sponges[17], early amphibian embryonic tissues[18] and bipinnaria larvae

---

[1] Neurogenesis and Organogenesis Group, RIKEN Center for Development Biology, Kobe 650-0047, Japan.

of starfish[19], in which tissue architectures that resemble the *in vivo* structures emerge. For instance, dissociated amphibian cells have been shown to form layered tissues that mimic the inside–outside order in the embryo: epidermis outside, and mesodermal and neural tissue inside[18]. The pattern formation in the reassembled tissue may be driven by complex processes, such as differential affinity between distinct elements (homophilic[20] compared with heterophilic[21] affinity, as well as mutual repulsion), spatial control of new attachment sites, active cell movements and cellular rearrangements[22,23].

Self-patterning

Complex tissue patterns can spontaneously appear from a homogeneous cell aggregate (or cell sheet) through spatiotemporal control of cell status by local rules and without external cues. A well-studied multicellular system for self-patterning is the cyclic-AMP-mediated wave pattern that is observed in the aggregation of the amoeba *Dictyostelium discoideum*[24]. Another classic example seen *in vitro* is the differentiation of plant calluses[25] into ordered structures, such as a stalk and roots, in the presence of the plant hormone auxin. Skin patterns in some fish[26], spacing of feathers[27] and hairs[28,29], and propagating waves of hair regeneration[30] are controlled tissue-autonomously by dynamic self-regulating rules. The emergence of polarized tissue patterns is seen in reaggregated animal cap cells treated with the growth factor activin[31] and has also been reported in three-dimensional (3D) mammalian cell cultures[32,33].

Self-driven morphogenesis

During organogenesis, complex but stereotyped tissue shapes develop through deformation, local growth and remodelling. In self-organization, multicellular morphogenesis is driven by intrinsic tissue mechanics and autonomously occurs in the absence of external forces or spatial constraints. For instance, spontaneous tubular formation from the epithelial cyst is seen in 3D culture of Madin–Darby canine kidney cells[34], in which cells are uniformly aggregated and cultured in homogeneous medium containing hepatocyte growth factor. Budding and branching morphogenesis is also seen in aggregate culture of salivary and mammary epithelial cells[35,36]. Self-driven morphogenesis should involve complex controls in the local generation of directed internal forces and dynamic modification of tissue stiffness and viscosity, as well as dynamic control of the cooperative response to mechanical force. These mechanisms also work in collective cell migration[37–39].

Tissue self-organization in 3D cell culture

Recently, remarkable progress has been made in understanding self-organization using cell cultures that allow *in vitro* recapitulation of organogenetic processes.

Embryonic stem cells

Unexpected levels of emergent self-organization have been observed for highly ordered structures, such as the fetal eye, using embryonic stem (ES) cells grown in culture[40–42] (Fig. 2a). During embryonic development, the retina first arises as a neuroepithelial vesicle that evaginates from the lateral wall of the rostral diencephalon, to form a structure called the optic vesicle. Subsequently, the distal portion of the optic vesicle invaginates to form a double-walled brandy-glass-shaped structure — the optic cup. The inner wall develops into the neural retina, containing photoreceptors and retinal neurons, whereas the outer wall differentiates into retinal pigment epithelium (RPE) — a supporting tissue for photoreceptor function and maintenance. This complex process of optic-cup formation has been successfully recapitulated in 3D culture of mouse and human ES cells[40,41] (Fig. 2a). A floating aggregate of 3,000 mouse ES cells has been cultured in medium suitable for retinal differentiation — it contained the extracellular matrix protein laminin. ES-cell-derived retinal epithelium first evaginates as vesicles and then undergoes invagination into the neural retina to form optic cups. Furthermore, this retinal epithelium, when isolated and cultured in suspension, gradually develops over the next 2 weeks into a multilayered structure with six different kinds of retinal cells, as observed in the postnatal eye[40]. This ES cell culture is started simply by reaggregation of dissociated ES cells, that is, with no particular prepattern. In addition, both optic-cup formation and multilayered neural retinal generation occur in homogeneous culture medium without external forces, positional cues or physical constraints.

Interestingly, optic-cup self-organization from ES cells is a size-sensitive phenomenon[40,41]. Retinal cell differentiation occurs even when the culture is started with a small aggregate of 300 mouse ES cells, but optic cups fail to form unless the number of cells reaches a critical mass (typically, 1,000–2,000 cells). Conversely, when a much larger aggregate (for example, more than 7,000 cells) is used, the size of the forming optic cups remains the same, but the number of forming cups increases.

Thus, under permissive conditions and with reasonable freedom (in this case, floating culture), the formation of optic cups and the multilayered neural retina occurs by self-organization, and is driven by an internal program that functions through local intercellular interactions (see go.nature.com/ii8n7a for a four-dimensional (4D) demonstration of this morphogenesis). A similar self-organizing phenomenon, albeit less complete, has been reported for multilayered cerebral cortex formation from mouse and human ES cells and from induced pluripotent stem (iPS) cells[42,43] (Fig. 2a).

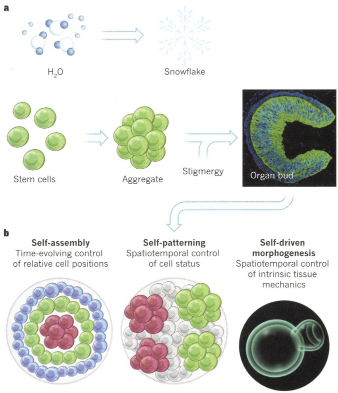

a
H_2O
Snowflake
Stem cells
Aggregate
Stigmergy
Organ bud
b
Self-assembly
Time-evolving control of relative cell positions
Self-patterning
Spatiotemporal control of cell status
Self-driven morphogenesis
Spatiotemporal control of intrinsic tissue mechanics

Figure 1 | Three basic mechanisms of tissue self-organization. **a**, Fundamental difference between self-organization in physics and in multicellular systems. Compared with self-organization in biological systems (shown here by the formation of an organ bud), whose elements change and local rules of interaction evolve spatiotemporally, the elements and local rules in physics and chemistry systems (shown here as the formation of a snowflake) are constant. Biological systems are also influenced by stigmergy — the history dependency of a system; in this case stigmergy affects the local rules of interaction at multiple steps after a cell aggregate forms. **b**, In biological systems multicellular self-organization involves a combination of self-assembly, self-patterning and self-driven morphogenesis. Self-assembly involves the time-evolving control of cell positions relative to each other, such as in the formation of a layered pattern in tissue development. Self-patterning is the spatiotemporal control of cell status, so that cells acquire heterogeneous properties in a region-specific manner from a homogeneous cell population. Self-driven morphogenesis is the spatiotemporal control of intrinsic tissue mechanics, shown here as a representation of deformation of retinal epithelium.

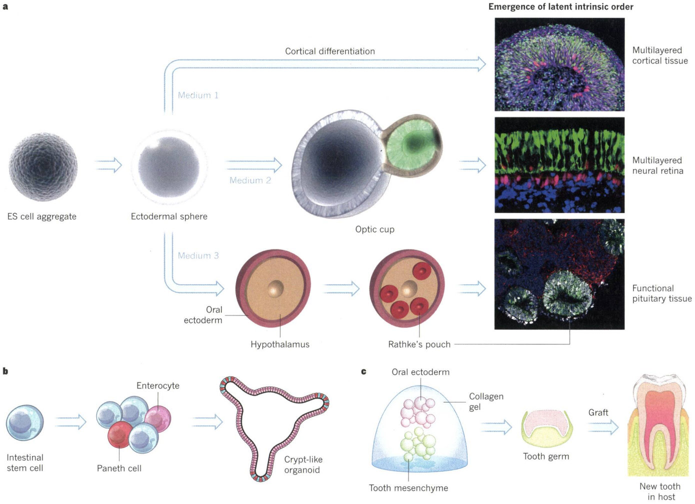


Figure 2 | Examples of tissue self-organization in three-dimensional cell culture. **a**, Embryonic stem (ES) cell culture with three-dimensional aggregates for self-organization, depending on the medium, into stratified cortical tissue (top), and the optic cup and the multilayered neural retina (middle) from homogenous progenitors. Other tissues, however, require interaction with different tissues in order to self-organize. For example, Rathke's pouch (bottom), from which pituitary tissue develops, forms from oral ectoderm under the influence of hypothalamic neuroectoderm. **b**, Self-formation of crypt-like organoids from a single-cell culture of intestinal stem cells that self-renew and produce cells such as Paneth cells and enterocytes, creating a niche in which the organoid can develop. **c**, Teeth develop by conjugation of 3D reaggregates of oral ectoderm and tooth mesenchyme cultured in collagen gel to generate a tooth-germ structure, which can grow into a tooth after transplantation into the host, such as mouse alveolar bone.

Optic-cup and cortex self-organization occurs from homogeneous progenitors, but many other cases of organogenesis require local interactions between different types of tissues in the initial phase. Such tissue–tissue interactions have been successfully recapitulated for development of the adenohypophysis (the hormone-producing portion of the pituitary gland) in 3D ES cell culture[44] (Fig. 2a) (see go.nature.com/phgmow for a demonstration of pituitary self-organization). During early development, Rathke's pouch (the adenohypophysis anlage) arises from oral ectoderm under the influence of the overlying hypothalamic neuroectoderm. When oral ectoderm and hypothalamic neuroectoderm are induced to differentiate together in a mouse ES cell aggregate under optimal conditions, they spontaneously produce the outer and inner layers of the aggregate, respectively. Then, multiple pituitary placodes self-form through the layers' interactions with each other and invaginate to develop Rathke's pouches, which subsequently generate functional pituitary endocrine cells that are suitable for transplantation[44].

Adult stem cells for epithelial tissue

A special micro-environment, or niche, is generally required for the growth and maintenance of adult stem cells. Our understanding of the relationship between tissue stem cells and their niche has been greatly advanced by recent *in vitro* studies that used a 3D culture of isolated intestinal stem cells. A gut 'organoid' that resembles the intestinal crypt and gut epithelial tissue can form clonally from an individual intestinal stem cell (therefore, without pre-patterning) in the presence of homogeneous extrinsic signals, including R-spondin proteins (which activate the Wnt-signalling pathway)[45] (Fig. 2b). Under these Wnt-enhanced conditions, crypt-base columnar cells (intestinal stem cells that express the R-spondin receptor Lgr5) self-renew and also produce other gut epithelial cells, including Paneth cells. The Paneth cells generated subsequently create a niche for intestinal stem cells, forming a self-sustaining system for stem cells in the 3D colony[46]. Crypt-like protrusions with tubular morphology form in a self-driven manner, whereas the epithelium of the cyst is composed of differentiated gut epithelial cells such as enterocytes. A similar self-organizing formation of gut organoids has been reported with pluripotent stem cells[47].

Primary culture cells

Recombination culture of primary culture cells can also exhibit a substantial level of emergence of highly ordered structures. For instance, teeth and hair are appendages that develop from oral and surface ectoderm, respectively, under the inductive influences of underlying local mesenchyme. The early developmental processes of these appendages have been successfully recapitulated in recent recombination studies using reaggregates[48–51]. In the case of teeth (Fig. 2c), the 3D structure of tooth germ self-forms *in vitro* from an aggregate of embryonic oral ectodermal cells in conjunction with a local mesenchymal aggregate grown in collagen gel[48,49]. The tooth germ develops into a fully grown tooth after grafting into alveolar bone[48]. Similarly, a fully functional hair follicle can self-form in a conjugation culture of embryonic surface ectodermal cells and dermal mesenchymal cells (or adult hair-bulge stem cells and papilla mesenchymal cells)[51]. In the case of self-formation from primary culture cells, however, the exact contribution of self-patterning to structural emergence requires careful consideration, given that substantial heterogeneity may pre-exist in the cell aggregates.

In vivo roles of self-organization

Previous studies have also shown that ectopic formation of complex structures is observed *in vivo* under certain circumstances. For instance, ectopic eyes are induced by *Pax6* misexpression in the fruitfly *Drosophila melanogaster*[52], whereas additional limb development from the flank occurs when fibroblast growth factors (FGFs) are applied locally in chick embryos[53], demonstrating that the internal programs for robustly autonomous organogenesis could be initiated by simple local triggers. These studies suggest that self-developing mechanisms that are dependent on local interactions work in various aspects of organ bud formation.

To investigate the *in vivo* and *in vitro* roles of self-organization more strictly, vertebrate eye development may be a good testing ground. A high level of robustness is needed for the *in vivo* generation of a precisely organized eye; the eyeball is of no use without an accurate arrangement of the lens and retinal axes. This robustness is presumably achieved by linking multiple principles. For instance, optic-cup invagination could be regulated more precisely in space and time when both optic-cup self-organization (internal rules)[40,41] and spatial constraints owing to lens invagination and cornea formation (external influences)[54,55] are cooperatively coupled. Interestingly, a similar idea was mentioned by the German embryologist Hans Spemann in his 1901 study on lens induction[56,57], "the complex apparatus of the vertebrate eye originates by a sequence of developmental processes which … are interlocked spatially and temporally".

Mechanisms for self-patterning

An important question is how the self-organization of pattern and shape can occur without prepatterns or external instructions, as seen in ES cell aggregate culture. In theory, to generate pattern from homogeneous cell populations, self-patterning should involve some sort of symmetry breaking[58], which is generally driven by stochastic processes that are stabilized by programmed intercellular interactions.

A role for the reaction–diffusion mechanism in many aspects of self-patterning has been suggested. In 1952, the British mathematician Alan Turing proposed spontaneous formation of a repeated or single 'stationary wave' pattern[59] (known as Turing pattern) from a non-patterned field, with a simple logic of mutual interactions that can be expressed by as few as two partial differential equations. In this form of interaction, the activator of a given pattern activates itself (auto-activation) and its inhibitor (for negative feedback) with different kinetics. Both the activator and inhibitor (or their effects) also act on surrounding elements in a distance-sensitive fashion, whereas the effect of the inhibitor should spread more efficiently than that of the activator (that is, short-range positive feedback with long-range negative feedback)[60,61]. The nature of the repeated pattern is dependent on the strength, delay and diffusion of auto-activation and feedback as well as the boundary conditions of the field. Importantly, in rapidly growing systems, such as a stem-cell aggregate or an early embryo, these parameters and the boundary conditions are continually changing. In these cases, the pattern formed by the reaction–diffusion mechanism may not easily settle, and its stable formation may be facilitated by coupling with some other mechanisms, such as epigenetic changes for memory in cells (typically, by differentiation).

The reaction–diffusion mechanism has been shown to function in many different aspects of body pattern formation. For example, a recent report has demonstrated that the generation of the regularly spaced transverse ridges of the palate is driven by a Turing-type reaction–diffusion system that involves FGF and the signalling protein Shh as components of an activator–inhibitor pair[62]. In addition to these relatively stationary patterns, feedback systems such as the reaction–diffusion mechanism can also generate oscillatory patterns and travelling waves. For instance, such temporal and spatial periodicity, which is driven by an internal clock, has an important role in tissue-autonomous segmentation in the paraxial mesoderm and in somitogenesis[63,64].

Another common mechanism of symmetry breaking in tissue concerns bistability in the regulatory network[58,65,66]. This can occur when state A and state B inhibit each other, while each state activates itself. Under these bistable conditions, any state between them is unstable and tends to flow into the local minimum created by state A or B. This mechanism is suitable for producing a clear polarity and a sharp boundary of domains, and is also known to create a certain level of hysteresis (for example, creating a toggle switch[66]), which can also contribute to symmetry breaking. In multidimensional regulation, bifurcation could be more complex and may involve multiple metastable local minima and saddle points. In rapidly growing systems, the bistability-based pattern may also require other mechanisms for stabilization such as a gradient of other patterning signals that spatially bias either state.

For instance, in optic-cup self-organization, the determination of the neural retina compared with RPE seems to involve bistable states of Wnt signalling, which promotes differentiation of RPE[67] (Fig. 3). Biased activation of Wnt signalling in RPE, once it happens, further enhances Wnt signalling by activating the expression of Wnt ligands in RPE and by inducing the transcription factor Mitf, which is implicated in Wnt signal facilitation[68] (Fig. 3b). Attenuated Wnt signalling in the neural retina makes the tissue receive lower strength Wnt signals by expressing soluble Wnt inhibitors[40] such as Dkk1 and Dkk3 and by suppressing Wnt-ligand expression[69] through Six3. Once the neural retina is induced, stabilization of the neural retina versus RPE pattern in the optic cup may be reinforced by FGF ligands[70], which are expressed by the neural retina and promote its differentiation at the cost of RPE differentiation.

Asymmetric cell division and planar cell polarity[58] can be major

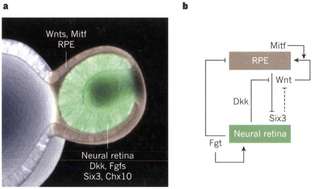


Figure 3 | Bistable signal interactions in regional patterning of the optic cup. Retinal pigment epithelium (RPE) and the neural retina inhibit each other's development through Wnt proteins — which promote RPE differentiation at the cost of neural retina differentiation — and Dkk (a diffusible Wnt antagonist). RPE reinforces its own fate by augmenting Wnt signals with increased expression of Wnt ligands and the transcription factor Mitf. Conversely, neural retina stabilizes its fate as a result of Wnt inhibition by the Wnt antagonist Dkk and the Wnt ligand expression suppressor Six3. Chx10 specifies neural retina fate, but is only indirectly related to Wnt signalling control. In addition, the neural retina produces FGFs, which promote the tissues differentiation at the cost of RPE differentiation. **a**, Representation of region-specific distribution of factors that influence differentiation of RPE and the neural retina in the self-organizing optic cup. **b**, A hypothetical relationship of mutual interactions between RPE and the neural retina in ES cell culture.

mechanisms that generate heterogeneity, particularly when coupled with certain boundary conditions. Indeed, in multicellular systems, there are some intrinsic boundary condition properties that can give non-equal spatial information to cells in the aggregate. For instance, a 3D aggregate inevitably has a superficial–central axis. Similarly, an epithelial sheet has an apical–basal polarity and a peripheral–central difference, unless it is spherical.

In real systems, these mechanisms do not work independently, but presumably function in concert with one another and with other mechanisms to transform a non-prepatterned cell mass into an organized tissue.

Mechanics of self-driven morphogenesis

Mechanical processes, which are central to tissue shaping, have begun to attract substantial attention from physicists and engineers, as well as experimental biologists in the field of cell and tissue biology. Intriguing topics in recent studies have included mechanochemical coupling and force sensing[71] through the hippo pathway[72–74], self-regulatory collective migration[37–39], tissue-specific control of stiffness and mechanical anisotropy[75,76], oscillatory tension generated by microfilament dynamics[77–79] and apoptosis-related force[80]. Aside from its direct role in tissue deformation, mechanical force may also provide a new signalling dimension[72–74] in the regulation of tissue self-organization, with a unique advantage in its rapid transmission across a large morphogenetic field.

A great advantage of a self-organizing system is that mechanical analyses can be simplified by excluding the influence of external forces or extrinsic physical constraints. For instance, in 3D mouse ES cell culture, optic cups form within a day from optic vesicles by self-driven deformation (invagination) of the retinal epithelium in four crucial phases of morphogenetic changes[40] (Fig. 4). In phase 1, the retinal epithelium forms a hemispheric vesicle (optic vesicle) by evagination, and contains a high level of activated myosin in both the distal and proximal portions. In phase 2, the distal epithelium, which is now committed to a neural retinal fate, becomes mechanically relaxed and flattened in curvature. In phase 3, the epithelial junction between the neural retina and RPE domains undergoes a strong apical constriction, making an acutely angled hinge domain. This causes a moderate bending of the neighbouring the neural retina epithelium inwards. In phase 4, the neural retina substantially expands in a tangential direction by rapid cell proliferation and invaginates to form an optic-cup shape. An important point here is that neural retina is mechanically flexible because of reduced myosin contractile activity, whereas RPE is stiff with a high concentration of active myosin (this actomyosin-dependent difference in mechanical stiffness has been experimentally confirmed by atomic-force-microscope-based elasticity analysis[40]). Surrounded by the stiff RPE, the expanding soft tissue of the neural retina cannot extend the hard outer shell, but instead has to deform itself by buckling inwards. On the basis of these observations, we have proposed a relaxation–expansion model for self-driven optic-cup morphogenesis[40,81]. This model does not depend on external forces or extrinsic physical constraints, but involves only three sequential rules of local mechanical changes: 1, relaxation of the neural retina; 2, apical constriction in the hinge domain; and 3, tangential expansion in the neural retina (Fig. 4). Computer simulation based on a spring-based vertex model has supported the feasibility of these three local rules for causing an invagination of the neural retina[40,81].

The key internal force that drives invagination during phase 4 is generated by the expansion of the neural retina, which is dependent on cell proliferation. Theoretically, such an expansion of tissue should generate compression as a result of the reaction force against the hard outer shell. This has been experimentally proven by 3D position-specific ablation analysis using a pinpointed intense femto-second multiphoton laser[40] (Fig. 4). When a cell in the invaginating neural retina was ablated, the gap was quickly filled by strong lateral compression. This compression disappeared when the tissue was pretreated with a mitotic inhibitor.

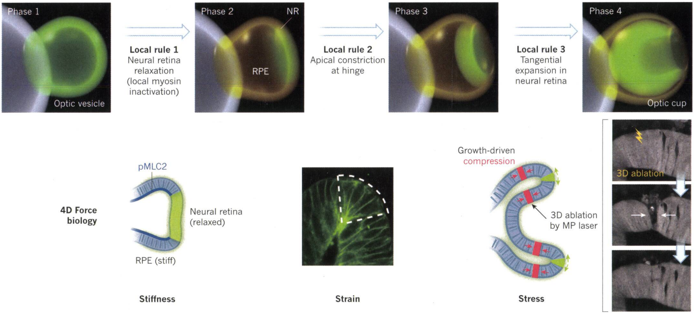


Figure 4 | Crucial controlling steps in self-driven optic-cup morphogenesis. Four phases of self-driven morphogenesis of the optic cup in embryonic stem cell culture. The retinal epithelium forms the optic cup by evagination (phase 1), the distal epithelium becomes relaxed and flattened (phase 2), the epithelial junction undergoes a strong apical constriction (phase 3), and the neural retina expands and invaginates to form an optic-cup shape (phase 4). Three local rules for the tissue mechanics involved have been elucidated[40,81]: rule 1, mechanical relaxation of the neural retina by local reduction of active myosin (loss of the phosphorylated regulatory subunit of myosin pMLC2 results in a loss of active myosin); rule 2, strong apical constriction of the hinge epithelium (making this portion of the epithelium wedge-shaped); and rule 3, rapid tangential expansion of the neural retina, which causes inward bucking of the tissue by generating compression. Compression in the neural retina was demonstrated by three-dimensional (3D) specific cell ablation through pinpointed irradiation with a multiphoton (MP) laser[40], in which the gap left by the abalated cell was quickly filled by strong lateral compression. How local mechanical rules for tissue stiffness, strain and stress are self-controlled in a phase-specific manner is a fundamental question for future studies in four-dimensional (4D) force biology.

A recent study of optic-cup formation from human ES cells has suggested that it also involves the three local rules[41]. Of note, the size of human ES-cell-derived optic cups is twice that of mouse ES-cell-derived ones (550–600 μm compared with 250–300 μm), which is consistent with the species-specific size difference between human and mouse optic cups *in vivo*[41]. The human neural retina is also thicker than that of the mouse (more than 150 μm), and therefore its invagination is mechanically more demanding. Probably because of this, human optic-cup formation involves an additional mechanism: self-driven eversion — a fourth local rule. In human (but not mouse) ES cell culture, the isolated neural retina spontaneously bends from its initially apically concave curvature into an apically convex curvature — in the same direction as it does during invagination. This eversion is associated with apically predominant volume expansion in neural retina epithelium, which is caused by apical deviation of progenitor cell bodies[41].

Thus, complex morphogenesis in optic-cup self-organization seems to be driven by a sequential execution of local mechanical rules, each of which is relatively simple.

Emerging technologies for multiplex dynamics

To understand the complex processes in tissue self-organization, it is necessary to analyse the dynamics of multicellular interactions in real time and space, and at the whole as well as the elemental level. A very promising technology that gives integrative multiscale information — *in toto* imaging[82–86] — can simultaneously record collective tissue morphogenesis (macroscopic) and individual behaviours of all cells (microscopic) in a system along the time span of the whole phenomenon. For instance, the four phases that occur during optic-cup morphogenesis have been elucidated by *in toto* imaging of the self-organizing culture[40]. Multichannel imaging provides multiplex information, that is, morphological information and functional aspects of cellular states. However, conventional optical devices have practical limitations for *in toto* imaging of tissue self-organization. For example, recording the entire optic-cup self-formation needs to cover several hundred micrometres along the *x*,*y*,*z* axes, but imaging beyond 100 μm in the *z* axis is generally difficult with visual light laser devices because of substantial diffraction in deep-tissue zones.

A new powerful technique for *in toto* imaging is light-sheet-based fluorescence microscopy (LSFM)[83–86], which uses selective light-sheet illumination in parallel to the *x* and *y* axes. Images of the *z* axis are obtained by shifting the position of the specimen swiftly. This optical system has important advantages: high brightness, fast speed and less phototoxicity. Another strong technology for *in toto* imaging is the multiphoton excitation (MPE) microscope[87], which uses an infrared femto-second pulse laser that is capable of penetrating deep into the tissue (typically more than 600–800 μm). In addition, MPE microscopy can be used for 3D laser ablation simultaneously with real-time imaging as already discussed[40]. Furthermore, MPE microscopy can use nonlinear optical imaging modes such as second and third harmonic generations, which do not require fluorescent labelling[88]. Compared with MPE, LSFM using visual light has limited light penetration into deep tissue (typically 200–300 μm). Recently, however, several groups have successfully combined LSFM technology with MPE[84–86] and applied it to 3D deep imaging of highly refractive fruitfly embryos using multidirectional observations.

Four-dimensional image bioinformatics[89] is also very important for the study of tissue dynamics. As the level of information obtained by *in toto* imaging has far exceeded that gained by manual evaluation, I see a great need for the development and sharing of high-performance and semi-automated image informatics tools for 4D multicellular analysis[89]. These tools include error-free multicell-tracking algorithms that can be applied to cell-dense environments and efficient extraction methods for morphological feature values[90]. The combination of *in toto* imaging and 4D image bioinformatics[83,88] is particularly important to elucidate history-dependent rules (stigmergy or interactions with a time lag), which are often otherwise latent.

A further important area of technical development is measurement of local mechanical properties, such as stiffness, surface tension, adhesion and expansion or retraction forces. Unlike advanced mechanical analysis in macroscopic tissues[91,92] and individual cells[71,93], methods for measurement at the mesoscopic scale (for example, domains in a optic cup) have not yet been sufficiently developed. Atomic-force microscope technology for measuring stiffness has been successfully applied to pin-pointed areas of cells[93] or monolayer tissues[40], whereas its application to 3D morphogenesis is still limited, partly because of the limited topological accessibility of its cantilever. Optical tweezers are excellent tools for protein-level mechanics, but the order of their force is generally too low for tissue deformation studies. By contrast, the force generated with ferromagnetic beads by strong electromagnetic coils or neodymium magnets is compatible with that in tissue mechanics, and they have been used to evaluate deforming or traction forces and tissue stiffness in the developing gut[75,76] and in cells undergoing collective migration[39].

A particularly challenging task is to invent the non-invasive methodology to spatiotemporally map stress and viscoelasticity within the tissue undergoing self-driven morphogenesis, in combination with strain mapping by *in toto* imaging. In addition, versatile methodology for distinguishing active and passive modes of deformation is needed. Successful development of tension probes, which was reported recently[94,95], may contribute to the advancement of these aspects of mechanical analysis. For instance, the visualization of stress distribution at focal adhesions with a fluorescence resonance energy transfer (FRET) probe that contains an elastic domain from spider-silk protein and is connected to the actin-filament-binding protein vinculin[94]. Other types of FRET probes have been applied to tension mapping in the microfilament system by insertion into α-actinin[95]. Another research group[96] has shown that stress at adherence junctions stretches a domain of α-catenin and transforms its protein structure into a vinculin-binding form[96], suggesting it has a potential use as a force-sensitive reporter. Further development of force measurement methodology is particularly crucial from the viewpoint of soft-matter physics, which is becoming an important approach towards understanding complex tissue mechanics[97].

Concept of cytosystems dynamics

Local interactions are multiplex in tissue self-organization; therefore, a change in an element's state (for example, differentiation) could simultaneously cause an alteration in interacting rules. This means that the 'epigenetic landscape' in such a complex system is not static (such as the classic canalization view described by the British embryologist C. H. Waddington) but rather has a self-evolving 'diachronic' nature. In this sense, its biology goes beyond simple structuralism and instead requires a post-structuralism viewpoint.

To understand such complex self-organizing mechanisms, conventional cell-level systems biology — which seeks to have a comprehensive analysis of elements and formulation of their interactions in the network (as beautifully demonstrated in a recent study of the bacterium *Mycoplasma genitalium*[98]) — is simply not applicable. Even for a relatively simple example, such as optic-cup formation, complete quantitative descriptions of the dynamic parameters of multicellular components and their interactions at biochemical, physical, genetic and epigenetic levels in space and time are completely unrealistic.

Therefore, for a mechanistic understanding of the dynamic nature of a self-developing system, we need to introduce a paradigm shift. In other words, instead of the reductionism-based, bottom-up, approach that describes all the details of a system, we need to deduce the core regulatory principles of systems dynamics from fragmented and discretized information. From this viewpoint, I discuss the prospect of a new direction in research for dynamic multicellular systems, which I propose as 'cytosystems dynamics'. This approach would comprise an interactive triad of complementary approaches: multiplex 4D measurement, multiscale mathematical modelling and multidirectional experimental reconstitution (Fig. 5a).

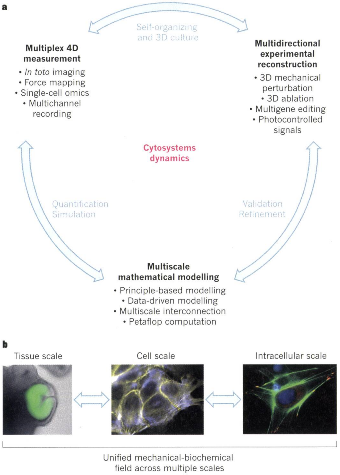

a
Self-organizing and 3D culture
Multiplex 4D measurement
• In toto imaging
• Force mapping
• Single-cell omics
• Multichannel recording
Multidirectional experimental reconstruction
• 3D mechanical perturbation
• 3D ablation
• Multigene editing
• Photocontrolled signals
Cytosystems dynamics
Quantification Simulation
Validation Refinement
Multiscale mathematical modelling
• Principle-based modelling
• Data-driven modelling
• Multiscale interconnection
• Petaflop computation
b
Tissue scale
Cell scale
Intracellular scale
Unified mechanical-biochemical field across multiple scales

Figure 5 | Cytosystems dynamics for multicellular emergence biology. **a**, To elucidate and manipulate 'acupoints' of dynamics (or core control molecules) in complex multicellular behaviours influenced by stigmergy, cytosystems dynamics would require the combination of three complementary approaches. These approaches are interconnected. Quantification of dynamics data by four-dimensional (4D) measurement is essential for mathematical modelling and data-driven simulation, whereas dynamic response analysis to perturbations in experimental reconstitution systems contributes to validation and refinement of mathematical models. For emergence biology in multicellular systems, a powerful tool for experimental reconstitution is self-organizing three-dimensional (3D) culture, which is suitable for *in toto* 4D imaging and experimental perturbation study. **b**, Multiscale interconnection in cytosystems dynamics for self-organization study. Intricate multiscale and stigmergetic interactions across tissue, cell and intracellular scales need to be analysed at multiple scales so that they can be linked together to elucidate complex cell behaviours. Among them, comprehensive understanding of the mechanical and biochemical fields (including mechanochemical coupling and feedback) is an important challenge.

For 4D measurement of multiplex data, a central player is the *in toto* imaging technology for simultaneous 4D recording of the whole (macroscopic) and the elements (microscopic) with wide, deep and high-resolution observations. Multichannel recording is used to obtain several categories of dynamic 4D information, including structural, biochemical, epigenetic (for example, cell differentiation) and mechanical data, at the same time and in the same population of cells. However, there is an apparent limitation on the multiplicity of simultaneous data acquisition; practically speaking, several simultaneous channels of information (up to ten or so) would be the technical maximum for live imaging even if different optical properties, such as wavelengths and polarization, were used. On the other hand, the rapidly advancing technology of high-content snapshots can provide supplementary information. For instance, a time series of genome-wide expression profiling data with single cells gives statistical information about varying states of the cells at each time point. For the interpretation of spatiotemporal dynamics of core regulatory principles, it is important to combine these high-context snapshot data bioinformatically with multiple sets of live imaging data that are rich in spatiotemporal information.

Another crucial challenge in obtaining an integrative understanding of a system is to use mathematical modelling for interpreting the dynamics of critical state transitions elucidated in 4D measurements. The modelling should involve bidirectional bottom-up and top-down approaches[5], that is, those based on the dynamics of known mechanisms in cell biology as well as those reinforced by data-driven modelling. To express complex collective cell behaviours in emergent self-organization, mathematical modelling must ultimately be done at multiple scales (Fig. 5b). For instance, to elucidate the emergent nonlinear nature of optic-cup morphogenesis, the tissue-scale model[81] needs to be linked not only with the cellular mechanics model but also with the models at the intracellular scales for cytoskeleton control and biochemical interactions. Invention of the methodology to link these multiscale models is wide open, and presumably needs to use approaches such as context-dependent conceptualization, contraction, coarse-graining and statistical estimation or optimization (for example, using Bayesian estimation), avoiding combinatorial explosion in computation. Large-scale parallel computation, now reaching the 10-petaflop level, will undoubtedly be a powerful tool for analysing cytosystems dynamics through multiscale computer simulations.

These models need to be validated and refined by experimental reconstitution of the system *in vitro*. In computer simulation of a complex multicellular system, just showing a similar phenotype *in silico* is not sufficient to support the model. It is important to demonstrate predictability of the model by simulating results for imaginary experiments with altered parameters that are experimentally controllable. The credibility of the model increases particularly if an expected simulation outcome with high nonlinearity is supported in reconstitution experiments. Notably, multicellular self-organization is full of nonlinear relationships and bifurcations, and *in vitro* reconstitution of the system (for example, using 3D culture of genetically modified ES cells) would allow us to test the model from multiple directions.

The three approaches discussed are actually interconnected. Regulatory parameters and boundary conditions in experimental reconstitution combined with 4D-controlled perturbations can be modified, for example, using photocontrollable signal modifiers[99], mechanical stimulation and laser ablation[40]. These dynamic data, when analysed by 4D imaging techniques, would provide crucial information for the refinement of multidimensional mathematical models. A study of chick embryos[100] is one example of pioneering efforts that combined quantitative imaging, modelling and experimentation. The study elegantly showed that the anterior–posterior body axis elongation is an emergent property that arises from the collective regulation of graded, random cell motion.

Basis for cytosystems dynamics

Finally, let us briefly discuss the basic aim and basis of the cytosystems dynamics approach in general. Often, this direction of research is inspired by a biological awe that we feel about collective behaviours or multicellular dynamics, such as observing the emergence of highly ordered phenotypes and their amazing robustness. When we observe individual cells in the system, their behaviours are complex and often seem stochastic owing to extrinsic and intrinsic fluctuations. Nonetheless, as a whole, cells work together to determine the pattern, structure, size and movement of their 'society' and the overall behaviour is reproducibly controlled.

Multicellular systems involve huge numbers of regulatory components, and the complexity of these in multidimensional networks is beyond comprehension. Such highly robust and regulative phenotypes could never be explained by strict control of each regulatory component in the system. Instead, it is inferred that robust interactions among a limited number of core control modules, occurring in a multiplex and multiscale manner, should have the predominant roles in the critical dynamics of the complex system, such as transitions in mechanical, biochemical and epigenetic states of locales.

In general, although each core control module must have some stability to form a self-sustaining local minimum, it must also be sensitive to dynamic state transitions (for example, bifurcation) upon crucial parameter alterations caused by the actions of other modules, acting as a tile in multidimensional 'dominoes'. At crucial regulatory steps (such as the transitions in the four phases during optic-cup morphogenesis), core control modules are expected to distinguish themselves as characteristic converging points in the switching dynamics of the network, regardless of the stochastic and complex nature of multicellular systems. Cytosystems dynamics seeks to identify these core control modules and elucidate their crucial interactions in executing the dynamic program correctly in time and space. For instance, such an approach should be necessary for explaining how three local rules for self-driven optic-cup morphogenesis are sequentially executed according to the internal program in a 4D context.

The concept of core control modules may be analogous to the idea of acupoints in acupuncture or 'tsubo' (a similar concept in Japan), for which stimulation exerts strong global effects on the autonomic nervous system. In this analogy, by identifying and steering the 'acupoints' of cytosystems dynamics, the hope is to manipulate complex multicellular behaviours such as self-organization, and perhaps even design them (for example their shape, size and functionality) for a new purpose. Although this is a long-term goal, I am rather optimistic about it. From recent iPS cell research, we know that a core control module that consists of only a few regulatory factors is sufficient to convert fibroblasts into pluripotent stem cells or cardiomyocytes. Likewise, mechanism-based control of tissue self-organization by a small number of acupoints should give us a new methodology of multicellular engineering that contributes to a wide range of biomedical applications. Given the advances in related technologies described in this Review, the time should be ripe for this new, challenging research field, for which I suggest the term emergence biology, with a goal to steer a multicellular population with complex behaviours on the basis of an understanding of its cytosystems dynamics. ■

Received 31 July; accepted 8 November 2012.

1. Nieuwkoop, P. D. Inductive interactions in early amphibian development and their general nature. *J. Embryol. Exp. Morphol.* **89,** S333–S347 (1985).
2. Camazine, S., Deneubourg, J. -L., Franks, N. R., Sneyd, J., Theraulaz, G. & Bonabeau, E. *Self-Organization in Biological Systems* (Princeton Univ. Press, 2001).
3. Yates, E. F., Garfinkle, A, Walter, D. O. & Yate, G. B. *Self-Organizing Systems: the Emergence of Order* (Plenum, 1987).
4. Chuong, C. M. & Richardson, M. K. Pattern formation today. *Int. J. Dev. Biol.* **53,** 653–658 (2009).
5. Saetzler, K., Sonnenschein, C. & Soto, A. M. Systems biology beyond networks: generating order from disorder through self-organization. *Semin. Cancer Biol.* **21,** 165–174 (2011).
6. Dobrescu, R. & Purcarea, V. I. Emergence, self-organization and morphogenesis in biological structures. *J. Med. Life* **4,** 82–90 (2011).
7. Whitesides, G. M. & Boncheva, M. Beyond molecules: self-assembly of mesoscopic and macroscopic components. *Proc. Natl Acad. Sci. USA.* **99,** 4769–4774 (2002).
8. Winfree, A. T. Spiral waves of chemical activity. *Science* **175,** 634–636 (1972).
9. Velarde, M. G. & Normand, C. Convection. *Sci. Am.* **243,** 92–108 (1980).
10. Forrest, S. B. & Haff, P. K. Mechanics of wind ripple stratigraphy. *Science* **255,** 1240–1243 (1992).
11. Kauffman, S. A. *The Origins of Order: Self-Organization and Selection in Evolution.* (Oxford Univ. Press, 1993).
12. Grassé, P. P. La reconstruction du nid et les coordinations inter-individuelles chez *Bellicositermes natalensis* et *Cubitermes* sp. La theorie de la stigmergie: Essai d'interpretation des termites. constructeurs. *Insectes Sociaux* **6,** 41–83 (1959).
13. Bonabeau, E. Stigmergy. *Artif. Life* **5,** 95–96 (1999).
14. Whitesides, G. M. & Grzybowski, B. Self-assembly at all scales. *Science* **295,** 2418–2421 (2002).
15. Caspar, D. L. & Klug, A. Physical principles in the construction of regular viruses. *Cold Spring Harb. Symp. Quant. Biol.* **27,** 1–24 (1962).
16. Gardner, M. K., Hunt, A. J., Goodson, H. V. & Odde, D. J. Microtubule assembly dynamics: new insights at the nanoscale. *Curr. Opin. Cell Biol.* **20,** 64–70 (2008).
17. Wilson, H. V. On some phenomena of coalescence and regeneration in sponges. *J. Exp. Zool.* **5,** 245–258 (1907).
18. Townes, P. L. & Holtfreter, J. Directed movements and selective adhesion of embryonic amphibian cells. *J. Exp. Zool.* **128,** 53–120 (1955).
19. Dan-Sohkawa, M., Yamanaka, H. & Watanabe, K. Reconstruction of bipinnaria larvae from dissociated embryonic cells of the starfish, *Asterina pectinifera. J. Embryol. Exp. Morphol.* **94,** 47–60 (1986).
20. Takeichi, M. Self-organization of animal tissues: cadherin-mediated processes. *Dev. Cell* **21,** 24–26 (2011).
21. Togashi, H. *et al.* Nectins establish a checkerboard-like cellular pattern in the auditory epithelium. *Science* **333,** 1144–1147 (2011).
In this article, the authors demonstrate how differential sets of heterophilic adhesion molecules can form chequerboard-like patterns in the epithelial sheet.
22. Mori, H., Gjorevski, N., Inman, J. L., Bissell, M. J., Nelson, C. M. Self-organization of engineered epithelial tubules by differential cellular motility. *Proc. Natl Acad. Sci. USA.* **106,** 14890–14895 (2009).
23. Tanner, K., Mori, H., Mroue, R., Bruni-Cardoso, A. & Bissell, M. J. Coherent angular motion in the establishment of multicellular architecture of glandular tissues. *Proc. Natl Acad. Sci. USA* **109,** 1973–1978 (2012).
24. Sawai, S., Thomason, P. A. & Cox, E. C. An autoregulatory circuit for long-range self-organization in *Dictyostelium* cell populations. *Nature* **433,** 323–326 (2005).
25. Thorpe, T. A. History of plant tissue culture. *Mol. Biotechnol.* **37,** 169–180 (2007).
26. Kondo, S. & Asai, R. A reaction–diffusion wave on the skin of the marine angelfish *Pomacanthus. Nature* **376,** 765–768 (1995).
This article is a pioneering paper that suggested Turing-type patterning occurs in stripe formation in the skin of fish.
27. Jiang, T. X., Jung, H. S., Widelitz, R. B. & Chuong, C. M. Self-organization of periodic patterns by dissociated feather mesenchymal cells and the regulation of size, number and spacing of primordia. *Development* **126,** 4997–5009 (1999).
28. Wang, Y., Badea, T. & Nathans, J. Order from disorder: self-organization in mammalian hair patterning. *Proc. Natl Acad. Sci. USA* **103,** 19800–19805 (2006).
29. Sick, S., Reinker, S., Timmer, J. & Schlake, T. WNT and DKK determine hair follicle spacing through a reaction–diffusion mechanism. *Science* **314,** 1447–1450 (2006).
30. Plikus, M. V. *et al.* Self-organizing and stochastic behaviors during the regeneration of hair stem cells. *Science* **332,** 586–589 (2011).
In this article the authors report how local interactions between hair follicles contribute to the global pattern of hair regeneration.
31. Green, J. B., Smith, J. C. & Gerhart, J. C. Slow emergence of a multithreshold response to activin requires cell-contact-dependent sharpening but not prepattern. *Development* **120,** 2271–2278 (1994).
32. ten Berge, D. *et al.* Wnt signaling mediates self-organization and axis formation in embryoid bodies. *Cell Stem Cell* **3,** 508–518 (2008).
33. Marikawa, Y., Tamashiro, D. A., Fujita, T. C. & Alarcon, V. B. Aggregated P19 mouse embryonal carcinoma cells as a simple *in vitro* model to study the molecular regulations of mesoderm formation and axial elongation morphogenesis. *Genesis* **47,** 103–106 (2009).
34. Montesano, R., Matsumoto, K., Nakamura, T. & Orci, L. Identification of a fibroblast-derived epithelial morphogen as hepatocyte growth factor. *Cell* **67,** 901–908 (1991).
35. Wei, C., Larsen, M., Hoffman, M. P. & Yamada, K. M. Self-organization and branching morphogenesis of primary salivary epithelial cells. *Tissue Eng.* **13,** 721–735 (2007).
36. Nelson, C. M., Vanduijn, M. M., Inman, J. L., Fletcher, D. A. & Bissell, M. J. Tissue geometry determines sites of mammary branching morphogenesis in organotypic cultures. *Science* **314,** 298–300 (2006).
37. Lecaudey, V., Cakan-Akdogan, G., Norton, W. H. & Gilmour, D. Dynamic Fgf signaling couples morphogenesis and migration in the zebrafish lateral line primordium. *Development* **135,** 2695–2705 (2008).
This paper describes the mechanism by which FGF signals drive morphogenesis in lateral-line organs during their collective migration.
38. Friedl, P. & Gilmour, D. Collective cell migration in morphogenesis, regeneration and cancer. *Nature Rev. Mol. Cell Biol.* **10,** 445–457 (2009).
39. Weber, G. F., Bjerke, M. A. & DeSimone, D. W. A mechanoresponsive cadherin-keratin complex directs polarized protrusive behavior and collective cell migration. *Dev. Cell* **22,** 104–115 (2012).
40. Eiraku, E. *et al.* Self-organizing optic-cup morphogenesis in three-dimensional culture. *Nature* **472,** 51–56 (2011).
This paper describes the how the optic cup self-organization mechanism is driven by local rules in ES cell culture.
41. Nakano, T. *et al.* Self-formation of optic cups and storable stratified neural retina from human ES cells. *Cell Stem Cell* **10,** 771–785 (2012).
42. Eiraku, M. *et al.* Self-organized formation of polarized cortical tissues from ESCs and its active manipulation by extrinsic signals. *Cell Stem Cell* **3,** 519–532 (2008).
This paper describes spontaneous layer formation in ES-cell-derived cortical tissues.
43. Mariani, J. *et al.* Modeling human cortical development *in vitro* using induced pluripotent stem cells. *Proc. Natl Acad. Sci. USA* **109,** 12770–12775 (2012).
44. Suga, H. *et al.* Self-formation of functional adenohypophysis in three-dimensional culture. *Nature* **480,** 57–62 (2011).
45. Sato, T. *et al.* Single Lgr5 stem cells build crypt–villus structures *in vitro* without a mesenchymal niche. *Nature* **459,** 262–265 (2009).
46. Sato. T. *et al.* Paneth cells constitute the niche for Lgr5 stem cells in intestinal crypts. *Nature* **469,** 415–418 (2011).

47. Spence, J. R., *et al.* Directed differentiation of human pluripotent stem cells into intestinal tissue *in vitro*. *Nature* **470,** 105–109 (2011).
48. Ikeda, E. *et al.* Fully functional bioengineered tooth replacement as an organ replacement therapy. *Proc. Natl Acad. Sci. USA* **106,** 13475–13480 (2009).
49. Nakao, K. *et al.* The development of a bioengineered organ germ method. *Nature Methods* **4,** 227–230 (2007).
50. Zheng, Y. *et al.* Organogenesis from dissociated cells: generation of mature cycling hair follicles from skin-derived cells. *J. Invest. Dermatol.* **124,** 867–876 (2005).
51. Toyoshima, K. E. *et al.* Fully functional hair follicle regeneration through the rearrangement of stem cells and their niches. *Nature Commun.* **3,** 784 (2012).
52. Halder, G., Callaerts, P. & Gehring, W. J. Induction of ectopic eyes by targeted expression of the eyeless gene in *Drosophila*. *Science* **267,** 1788–1792 (1995).
53. Cohn, M. J., Izpisúa-Belmonte, J. C., Abud, H., Heath, J. K. & Tickle, C. Fibroblast growth factors induce additional limb development from the flank of chick embryos. *Cell* **80,** 739–746 (1995).
54. Adler, R. & Canto-Soler, M. V. Molecular mechanisms of optic vesicle development: complexities, ambiguities and controversies. *Dev. Biol.* **305,** 1–13 (2007).
55. Fuhrmann, S. Eye morphogenesis and patterning of the optic vesicle. *Curr. Top. Dev. Biol.* **93,** 61–84 (2010).
56. Spemann, H. Über korrelationen in der Entwicklung des Auges. *Verh. Anat. Ges.* **15,** 61–79 (1901).
57. Hamburger, V. *The Heritage of Experimental Embryology* (Oxford Univ. Press, 1988).
58. Li, R. & Bowerman, B. *Symmetry Breaking in Biology* (Cold Spring Harbor Press 2010).
59. Turing, A. The chemical basis of morphogenesis. *Phil. Trans. R. Soc. Lond. B* **237,** 37–72 (1952).
60. Meinhardt, H. & Gierer, A. Pattern formation by local self-activation and lateral inhibition. *BioEssays* **22,** 753–760 (2000).
61. Kondo, S. & Miura, T. Reaction–diffusion model as a framework for understanding biological pattern formation. *Science* **329,** 1616–1620 (2010).
62. Economou, A. D. *et al.* Periodic stripe formation by a Turing mechanism operating at growth zones in the mammalian palate. *Nature Genet.* **44,** 348–351 (2012).
63. Palmeirim, I., Henrique, D., Ish-Horowicz, D. & Pourquié, O. Avian hairy gene expression identifies a molecular clock linked to vertebrate segmentation and somitogenesis. *Cell* **91,** 639–648 (1997).
    This paper demonstrates that vertebrate segmentation occurs in a tissue-autonomous manner that is dependent on its molecular clock.
64. Palmeirim, I., Dubrulle, J., Henrique, D., Ish-Horowicz, D. & Pourquié, O. Uncoupling segmentation and somitogenesis in the chick presomitic mesoderm. *Dev. Genet.* **23,** 77–85 (1998).
65. Ferrell, J. E. Jr. Bistability, bifurcations, and Waddington's epigenetic landscape. *Curr. Biol.* **22,** R458–R466 (2012).
66. Gardner, T. S., Cantor, C. R. & Collins, J. J. Construction of a genetic toggle switch in *Escherichia coli*. *Nature* **403,** 339–342 (2000).
67. Fuhrmann, S. Wnt signaling in eye organogenesis. *Organogenesis* **4,** 60–67 (2008).
68. Yasumoto, K. *et al.* Microphthalmia-associated transcription factor interacts with LEF-1, a mediator of Wnt signaling. *EMBO J.* **21,** 2703–2714 (2002).
69. Liu, W., Lagutin, O., Swindell, E., Jamrich, M. & Oliver, G. Neuroretina specification in mouse embryos requires Six3-mediated suppression of Wnt8b in the anterior neural plate. *J. Clin. Invest.* **120,** 3568–3577 (2010).
70. Zhao, S. *et al.* Patterning the optic neuroepithelium by FGF signaling and Ras activation. *Development* **128,** 5051–5060 (2001).
71. Hoffman, B. D., Grashoff, C. & Schwartz, M. A. Dynamic molecular processes mediate cellular mechanotransduction. *Nature* **475,** 316–323 (2011).
72. Dupont, S. *et al.* Role of YAP/TAZ in mechanotransduction. *Nature* **474,** 179–183 (2011).
73. Wada, K., Itoga, K., Okano, T., Yonemura, S. & Sasaki, H. Hippo pathway regulation by cell morphology and stress fibers. *Development* **138,** 3907–3914 (2011).
    References 72 and 73 together demonstrate crucial roles for hippo-related pathways in cellular mechanotransduction, leading to cellular growth and differentiation.
74. Sansores-Garcia, L. *et al.* Modulating F-actin organization induces organ growth by affecting the Hippo pathway. *EMBO J.* **30,** 2325–2335 (2011).
75. Taniguchi, K. *et al.* Chirality in planar cell shape contributes to left-right asymmetric epithelial morphogenesis. *Science* **333,** 339–341 (2011).
76. Savin, T. *et al.* On the growth and form of the gut. *Nature* **476,** 57–62 (2011).
    This paper is a pioneering report of morphogenetic tissue mechanics, using quantitative measurement of local tissue viscoelasticity.
77. Martin, A. C., Kaschube, M. & Wieschaus, E. F. Pulsed contractions of an actin–myosin network drive apical constriction. *Nature* **457,** 495–499 (2009).
78. Rauzi, M., Lenne, P. F. & Lecuit, T. Planar polarized actomyosin contractile flows control epithelial junction remodelling. *Nature* **468,** 1110–1114 (2010).
79. He, L., Wang, X., Tang, H. L. & Montell, D. J. Tissue elongation requires oscillating contractions of a basal actomyosin network. *Nature Cell Biol.* **12,** 1133–1142 (2010).
80. Toyama, Y., Peralta, X. G., Wells, A. R., Kiehart, D. P. & Edwards, G. S. Apoptotic force and tissue dynamics during *Drosophila* embryogenesis. *Science* **321,** 1683–1686 (2008).
81. Eiraku, M., Adachi, T. & Sasai, Y. Relaxation-expansion model for self-driven retinal morphogenesis. *BioEssays* **34,** 17–25 (2012).
82. Megason, S. G. & Fraser, S. E. Digitizing life at the level of the cell: high-performance laser-scanning microscopy and image analysis for *in toto* imaging of development. *Mech. Dev.* **120,** 1407–1420 (2003).
83. Keller, P. J., Schmidt, A. D., Wittbrodt, J. & Stelzer, E. H. Reconstruction of zebrafish early embryonic development by scanned light sheet microscopy. *Science* **322,** 1065–1069 (2008).
    This article is a pioneering study of *in toto* imaging for cell-lineage tracing.
84. Truong, T. V., Supatto, W., Koos, D. S., Choi, J. M. & Fraser, S. E. Deep and fast live imaging with two-photon scanned light-sheet microscopy. *Nature Methods* **8,** 757–760 (2011).
85. Tomer, R., Khairy, K., Amat, F. & Keller, P. J. Quantitative high-speed imaging of entire developing embryos with simultaneous multiview light-sheet microscopy. *Nature Methods* **9,** 755–763 (2012).
86. Krzic, U., Gunther, S., Saunders, T. E., Streichan, S. J. & Hufnagel, L. Multiview light-sheet microscope for rapid *in toto* imaging. *Nature Methods* **9,** 730–733 (2012).
87. Helmchen, F. & Denk, W. Deep tissue two-photon microscopy. *Nature Methods* **2,** 932–940 (2005).
88. Olivier, N. *et al.* Cell lineage reconstruction of early zebrafish embryos using label-free nonlinear microscopy. *Science* **329,** 967–971 (2010).
89. Eliceiri, K. W. *et al.* Biological imaging software tools. *Nature Methods* **9,** 697–710 (2012).
90. Zhong, Q., Busetto, A. G., Fededa, J. P., Buhmann, J. M. & Gerlich, D. W. Unsupervised modelling of cell morphology dynamics for time-lapse microscopy. *Nature Methods* **9,** 711–713 (2012).
91. Moore, S. W., Keller, R. E. & Koehl, M. A. The dorsal involuting marginal zone stiffens anisotropically during its convergent extension in the gastrula of *Xenopus laevis*. *Development* **121,** 3131–3140 (1995).
92. Davidson, L. A. Embryo mechanics: balancing force production with elastic resistance during morphogenesis. *Curr. Top. Dev. Biol.* **95,** 215–241 (2011).
93. Krieg, M. *et al.* Tensile forces govern germ-layer organization in zebrafish. *Nature Cell Biol.* **10,** 429–436 (2008).
94. Grashoff, C. *et al.* Measuring mechanical tension across vinculin reveals regulation of focal adhesion dynamics. *Nature* **466,** 263–266 (2010).
95. Meng, F. & Sachs, F. Orientation-based FRET sensor for real-time imaging of cellular forces. *J. Cell Sci.* **125,** 743–750 (2012).
96. Yonemura, S., Wada, Y., Watanabe, T., Nagafuchi, A. & Shibata, M. α-Catenin as a tension transducer that induces adherens junction development. *Nature Cell Biol.* **12,** 533–542 (2010).
97. Hannezo, E., Prost, J. & Joanny, J. F. Instabilities of monolayered epithelia: shape and structure of villi and crypts. *Phys. Rev. Lett.* **107,** 078104 (2011).
98. Karr, J. R. *et al.* A whole-cell computational model predicts phenotype from genotype. *Cell* **150,** 389–401 (2012).
99. Wu, Y. I. *et al.* A genetically encoded photoactivatable Rac controls the motility of living cells. *Nature* **461,** 104–108 (2009).
100. Bénazéraf, B. *et al.* A random cell motility gradient downstream of FGF controls elongation of an amniote embryo. *Nature* **466,** 248–252 (2010).

Acknowledgements I would like to thank my collaborators, especially M. Eiraku, N. Takata, K. Muguruma and T. Adachi, for stimulating strategic discussion on multicellular self-organization.

Author information Reprints and permission information is available at www.nature.com/reprints. The author declares no competing financial interests. Readers are welcome to comment on the online version of this article at go.nature.com/lu2gjc. Correspondence should be addressed to Y.S. (yoshikisasai@cdb.riken.jp).

REVIEW

doi:10.1038/nature11860

Activity-dependent neuronal signalling and autism spectrum disorder

Daniel H. Ebert[1,2] & Michael E. Greenberg[1]

Neuronal activity induces the post-translational modification of synaptic molecules, promotes localized protein synthesis within dendrites and activates gene transcription, thereby regulating synaptic function and allowing neuronal circuits to respond dynamically to experience. Evidence indicates that many of the genes that are mutated in autism spectrum disorder are crucial components of the activity-dependent signalling networks that regulate synapse development and plasticity. Dysregulation of activity-dependent signalling pathways in neurons may, therefore, have a key role in the aetiology of autism spectrum disorder.

About 1 in 100 children display signs and symptoms that lead to a diagnosis of autism spectrum disorder (ASD)[1]. This debilitating developmental disorder is characterized by impairments in social interaction and communication, and by restricted, repetitive and stereotyped behaviour and interests. In addition, individuals with ASD often have a seizure disorder and intellectual disability[2]. Most of the features of ASD manifest in the first few years of life, at the time of brain development when sensory experience is modifying excitatory synapse maturation and elimination, and promoting the development of inhibitory synapses[3]. This has led to the hypothesis that ASD may be due to a disruption of the normal process of experience-dependent synaptic development, resulting in an imbalance between excitation and inhibition in the developing brain.

Research has shown unequivocally that ASD is largely a genetic disorder[4]. Advances in human genetics and next-generation sequencing have identified a stunning number of newly arising mutations that are linked to ASD. In some cases, the mutations are limited to a specific gene and result in syndromic neurodevelopmental disorders with highly penetrant features of autism. These include fragile X syndrome, tuberous sclerosis complex, Angelman syndrome, Timothy syndrome and Rett syndrome[2]. In addition, an impressive array of human genetic studies, including studies of families with several children with ASD and more recent exome sequencing of trios of individuals (mother, father and child with ASD), has surprisingly shown that ASD is often due to newly arising gene copy number variants (CNVs) — such as a deletion or duplication of a region of a chromosome — or a rare mutation that arises in the germ cell, particularly in the sperm of older fathers[5–10]. In cases with CNVs, ASD is hypothesized to be a result of the increased or decreased expression of one or several genes that lie within the region of the genome in which the CNV mutation resides.

A future challenge is to make biological sense of the large number, and diversity, of genes that are associated with ASD. Identifying convergent molecular pathways in which multiple candidate genes are involved may be an effective way to advance our understanding of the underlying molecular basis of ASD, as well as a way to develop treatments for a broad range of forms of this disorder. Recent studies suggest that a convergent molecular pathway dysregulated in ASD is the signalling network that controls synapse development and function. An interesting feature of this signalling network is that it is composed of many proteins for which expression or function is regulated by neuronal activity, suggesting that one cause of ASD may be the dysregulation of neuronal activity-dependent synapse development and function.

A cardinal feature of human brain development is that sensory, cognitive and emotional experiences shape synapse and neural-circuit development. Neuronal activity triggers local changes at the synapse, altering the composition, shape and strength of the synapse. Synaptic stimulation both induces specific changes in messenger RNA translation near synapses and sends signals to the nucleus to induce gene transcription programs that control synaptic maturation and function. These neuronal activity-dependent pathways are crucial for learning and memory and for adaptive behavioural responses[11–13].

In this Review, we explore the hypothesis that the dysregulation of activity-dependent signalling networks that control synapse development and function may be an important component of the molecular basis of ASD. For instance, ASD may arise from dysregulation of activity-dependent signalling pathways locally at a synapse, including changes in the post-translational modifications of synaptic proteins and the local translation of mRNAs at synapses, as well as from the dysregulation of activity-dependent gene transcription. Alternative explanations for the molecular and cellular basis of ASD must also be considered, including the possibility that it is a result of dysregulation of synaptic function in general, impairment of neurotransmission (which secondarily may alter activity-dependent signalling) and defects in earlier steps in nervous-system development. To begin to evaluate these hypotheses, we review studies of the molecular function of the most studied genes that are associated with ASD, including those involved in syndromic disorders with penetrant autism spectrum features and rare mutations associated with ASD. The current weight of evidence does not conclusively demonstrate that dysregulation of activity-dependent neuronal signalling is the singular cause of ASD. Indeed, the impairment of different molecular pathways that are not mutually exclusive, probably contributes to the development of various subsets, or aspects, of ASD. However, as we review here, there is emerging evidence that genes and proteins associated with ASD are both regulated by and control activity-dependent pathways that modulate synaptic function. Many of the mutations associated with ASD lead to alterations in excitatory or inhibitory neurotransmission that disrupt activity-dependent signalling and activity-dependent synapse development, maturation and refinement. In addition, neuronal activity clearly regulates the function, localization and expression of many of the proteins that are associated with ASD. Taken together, these findings are

---

[1]Department of Neurobiology, Harvard Medical School, Boston, Massachusetts 02115, USA. [2]Department of Psychiatry, Massachusetts General Hospital, Harvard Medical School, Boston, Massachusetts 02114, USA.

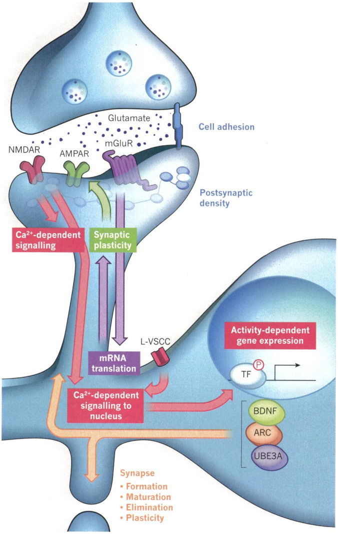


Figure 1 | Regulation of synaptic development and function by neuronal activity. Adhesion molecules and components of the postsynaptic density organize and regulate the formation of excitatory synapses on dendritic spines and can stimulate activity-dependent signalling networks within the postsynaptic neuron. At the synapse, the excitatory neurotransmitter glutamate can bind several glutamate receptors — the NMDA (*N*-methyl-D-aspartate) receptor (NMDAR), the AMPA (α-amino-3-hydroxy-5-methyl-4-isoxazole propionic acid) receptor (AMPAR) and the metabotropic glutamate receptor (mGluR). Signalling from stimulated mGluR regulates mRNA translation, which is required for long-lasting forms of synaptic plasticity. Modification and cell-surface expression of AMPA receptors underlies many aspects of synaptic plasticity. Stimulated NMDA receptors flux calcium, inducing calcium-dependent signalling networks, at the synapse, that regulate AMPA receptor function and actin reorganization. Calcium influx through NMDA receptors and L-type voltage-sensitive calcium channels (L-VSCCs) triggers calcium-dependent signalling to the nucleus, leading to the modification of transcriptional regulators and resulting in the induction of activity-dependent gene expression. Genes induced by neuronal activity (including *Bdnf*, *Arc* and *Ube3A*) function to regulate synapse formation, maturation, elimination and plasticity. P, phosphorylation; TF, transcription factor.

beginning to provide compelling evidence that disruption of the activity-dependent molecular programs that control synaptic function significantly contributes to the molecular basis of ASD.

Neuronal activity regulates synapses and circuits

Experience drives the neuronal activity that regulates synaptic development and plasticity, thereby altering the structure and function of neural circuits (Fig. 1). In a classic set of experiments done nearly 50 years ago[14], David Hubel and Torsten Wiesel discovered that occluding one eye during a crucial period of development disrupts the formation of ocular dominance columns in the visual cortex, demonstrating a central role for experience in the development of neural circuits. We now know from studies in hundreds of laboratories that synapses are stabilized, matured and eliminated in response to neuronal activity during postnatal development, and the molecular underpinnings of these processes have been significantly detailed[15].

The structure and the protein composition of excitatory synapses are crucial for synaptic function and the ability of the organism to respond to its environment. Synapses are highly specialized compartments within the neuron that comprise presynaptic and postsynaptic regions and a synaptic cleft for which function is regulated both spatially and temporally. Synaptic cell-adhesion molecules modulate specific steps in synapse formation and function, including the specialization of presynaptic and postsynaptic compartments that occurs during synapse development, the recruitment of molecules and vesicles into synapses, the modulation of *trans*-synaptic signalling, and the synaptic plasticity that is induced in response to sensory stimuli[16]. Excitatory synapses form mostly on dendritic spines — protrusions from dendrites that function as a finely tuned morphological and functional compartment. The postsynaptic density (PSD) forms at the centre of the dendritic spine head. The PSD contains structural proteins and signalling molecules, as well as receptors that respond to the excitatory neurotransmitter glutamate. PSD signalling proteins include the calcium-dependent kinases CaMKIIα and CaMKIIβ, and the scaffold protein PSD-95, which helps to organize the synapse, connect glutamate receptors to downstream signalling complexes and modulate synaptic plasticity. Other PSD scaffolding proteins include GKAP (also known as SAPAP), SHANK and HOMER proteins, which promote dendritic spine growth, regulate synaptic activity and anchor signalling complexes that mediate aspects of the synaptic response to sensory stimuli[17–19] (Fig. 2).

Experience leads to the release of neurotransmitters at specific synapses. Glutamate — the most common excitatory neurotransmitter — binds to the receptors for AMPA (α-amino-3-hydroxy-5-methyl-4-isoxazole propionic acid) and NMDA (*N*-methyl-D-aspartate). These ionotropic receptors then flux Na^+ and K^+, leading to membrane depolarization. AMPA receptors mediate fast excitatory neurotransmission and can be dynamically inserted and removed from the synaptic membrane, thus affecting the strength of the postsynaptic response to glutamate released at the synapse. NMDA receptors open up and conduct current only when both the postsynaptic membrane is depolarized and glutamate binds the receptor, providing coincidence detection and a signal that induces forms of synaptic plasticity[20]. Activated NMDA receptors flux calcium, thereby inducing multiple calcium-dependent signalling pathways within dendrites, including CaMKs and the Ras–mitogen-activated protein kinase (MAPK) pathway[21], which alter the synapse locally and stimulate gene transcription in the nucleus. For instance, activated CaMKII phosphorylates AMPA receptors, altering their conductivity, and triggering reorganization of the actin cytoskeleton to regulate dendritic growth[15]. Neuronal activity also induces long-lasting changes in the efficacy of synaptic transmission, including long-term potentiation (LTP) and long-term depression (LTD), which involve the strengthening or weakening of synaptic transmission, respectively, and are thought to be molecular correlates of learning and memory[11,20,22]. Depending on the nature of the stimulus, different forms of synaptic plasticity are induced at synapses. LTP or LTD are mediated, at least in part, through the insertion or removal, respectively, of AMPA receptors from the synaptic membrane, leading to alterations in synaptic strength (Fig. 2).

Glutamate release at synapses can also induce longer-lasting forms of synaptic plasticity that are mediated by NMDA-receptor-dependent calcium influx and typically require activity-dependent protein synthesis, which is a consequence of the local mRNA translation within the dendrite or of gene transcription within the nucleus. In a similar manner, glutamate binding to metabotropic glutamate receptors (mGluR) induces protein-synthesis-dependent forms of synaptic plasticity by activating the mechanistic target of rapamycin (mTOR) and extracellular signal-regulated kinase (ERK) pathways[11,20,22,23].

Neuronal activity also regulates programs of gene transcription that affect synapses, including key steps in synaptic development and forms of synaptic plasticity[12,15]. Typically, activity-dependent gene transcription requires calcium-dependent signalling and, depending on the route and dynamics of calcium entry (either through NMDA receptors or through L-type voltage-sensitive calcium channels), can trigger distinct gene expression responses[24,25]. Neuronal activity-induced elevations in intracellular calcium lead to the activation of multiple signalling molecules, including CaMKII, CaMKIV, protein kinase A (PKA), MAPK pathways, ribosomal S6 kinase (RSK), mitogen-activated and stress-activated kinase (MSK), and the calcium-dependent phosphatase calcineurin, each of which phosphorylate or dephosphorylate transcriptional regulators within the nucleus. For instance, multiple stimuli induce the phosphorylation of the cyclic-AMP-dependant-responsive-element-binding protein (CREB) at Ser 133, which recruits the CREB-binding protein (CBP) to gene promoters. CBP catalyses the acetylation of histones, leading to changes in chromatin structure that facilitate the recruitment of RNA polymerase II and the activation of gene transcription programs that promote synapse development[12,15].

In addition to stimulating CREB-dependent gene transcription, neuronal activity-dependent calcium influx into neurons triggers the dephosphorylation of the transcription factor myocyte enhancer factor 2 (MEF2) by calcineurin, disrupting the association of MEF2 with histone deacetylases and leading to the recruitment of CBP and the stimulation of gene transcription[26]. Neuronal activity that functions through CREB, MEF2 and multiple other activity-regulated transcription factors — including SRF, Fos[27] and NPAS4 (ref. 28) — induces the transcription of the genes encoding proteins that function directly at synapses, including *Bdnf*, *Arc* and *Ube3A*[12,15]. For example, activity-regulated MEF2 (by activating *Arc*) suppresses the number of excitatory synapses[26,29], whereas CREB and NPAS4 (by activating *Bdnf* transcription) control the number of inhibitory synapses that form on excitatory neurons[28,30].

Many of the proteins that form the activity-dependent signalling network that controls gene transcription are implicated in disorders of cognition and behaviour, including ASD (Fig. 3). For example, mutations in the L-type voltage-sensitive calcium channels that flux calcium to initiate activity-dependent gene transcription are associated with Timothy syndrome, which is characterized by impairments in cardiac and immune function and has a significant association with ASD phenotypes[31]. Likewise, mutations in RSK2, CBP, the methyl-binding protein MECP2 and the ubiquitin ligase UBE3A are causes of Coffin–Lowry syndrome, Rubinstein–Taybi syndrome, Rett syndrome and Angelman syndrome, respectively[3]. The finding that many of the genes that are involved in activity-dependent gene transcription are associated with human disorders of cognitive function, including ASD, provides support for the hypothesis that dysregulation of activity-dependent signalling networks in general may contribute significantly to the synaptic dysfunction that occurs in ASD.

Genetic architecture of ASD

ASD is caused predominantly by genetic mutations[4]. Recent evidence suggests that *de novo* mutations that arise in the germ line are a significant cause of ASD, although it is important to keep in mind that the genetic aetiology of about 70% of the forms of ASD is unknown and that it remains perplexing why ASD is at least four times more prevalent in males than in females. Higher age of the father is positively correlated with a higher rate of *de novo* genetic mutations and an increased risk of ASD[9,10]. ASD is polygenic in aetiology; recent studies of *de novo* exonic mutations suggest the involvement of as many as 400 to 1,000 genetic loci in ASD[9]. The complexity of the genetic aetiology of ASD may complicate the understanding of its molecular mechanisms. However, as already suggested, many of the genes implicated in ASD could function in convergent molecular pathways, allowing for the formulation of specific mechanistic hypotheses that could explain significant portions of ASD.

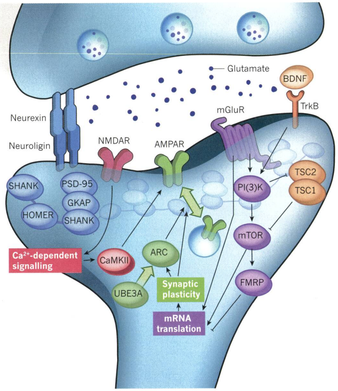


Figure 2 | Neuronal activity regulates mRNA translation and synaptic plasticity. At the excitatory synapse, the cell-adhesion molecules neurexin and neuroligin and structural proteins in the postsynaptic density, including SHANK proteins, regulate, and are regulated by, neuronal activity-dependent signalling networks. The excitatory neurotransmitter glutamate binds NMDA receptors (NMDARs), AMPA receptors (AMPARs) and mGluRs. AMPA receptors mediate fast excitatory neurotransmission. UBE3A degrades ARC, which can affect the trafficking of AMPA receptors. During plasticity, activation of NMDA receptors triggers calcium-dependent signalling at the synapse, stimulating CaMKII and modifying AMPA receptor function and actin reorganization. Activation of mGluR triggers several signalling cascades, including the phosphatidylinositol-3-OH kinase (PI(3)K)–AKT–mTOR pathway, Ras–mitogen-activated protein kinase (MAPK) pathway and PLC to regulate mRNA translation. FMRP inhibits mRNA translation, and mGluR signalling regulates FMRP activity. Growth factors, including BDNF, whose expression is induced by neuronal activity, bind receptor tyrosine kinases (including TrkB), which activate multiple signalling pathways, including the PI(3)K–AKT pathway. The PI(3)K–AKT pathway, when activated, leads to phosphorylation of the TSC1–TSC2 complex to control mTOR activity. Mutations in neurexins, neuroligins, SHANKs, GKAP, UBE3A, FMRP and TSC1–TSC2 are associated with ASD.

Rare genetic variations play a significant part in the aetiology of ASD. Several *de novo* CNVs that are associated with ASD have been detected multiple times, including CNVs at 16p11.2, 15q11–13, 22q11.2, deletions at NRXN1, and duplications at 7q11.23 (refs 5, 32). The fact that these CNVs have been detected in multiple individuals with ASD strongly suggests that mutations in these regions of the genome are contributing to ASD. More subtle, rare mutations in *NLGN4*, *NLGN3*, *NRXN1*, *SHANK2*, *SHANK3* and *CNTNAP2* (ref. 5) have also been identified in multiple individuals with ASD, suggesting that these mutations also contribute to these disorders. In addition, very recent exome sequencing of parent–offspring trios has identified *de novo* exonic mutations in *SCN2A*, *KATNAL2*, *CHD8*, *FOXP1*, *NTNG1*, *GRIN2B* and *LAMC3* (refs 7–9), extending further

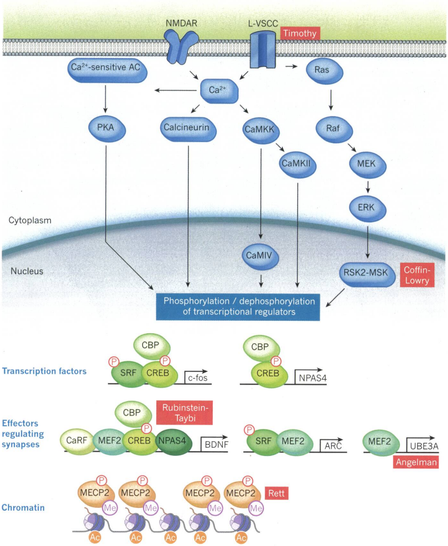


Figure 3 | Activity-dependent gene expression and ASD. Activation of NMDA receptors (NMDARs) and L-type voltage-sensitive calcium channels (L-VSCCs) triggers calcium influx and induction of calcium-dependent signalling molecules, including CaMKK (which activates the kinases CaMKII and CaMKIV), calcium-sensitive adenylyl cyclases (ACs) (which activate PKAs) and the calcium-dependent phosphatase calcineurin. In addition, the Ras–ERK–RSK/MSK pathway is activated upon stimulation of L-VSCCs. These signalling cascades phosphorylate or dephosphorylate transcriptional regulators in the nucleus, leading to the expression of transcription factors (shown as c-fos and NPAS4). These transcription factors contribute to the regulation of activity-dependent gene transcription and to the expression of effectors that function to modify synapses (shown as BDNF, ARC and UBE3A). In addition, changes in chromatin modifications (including histone acetylation), and the phosphorylation of chromatin-associated proteins (for example, MECP2), may modulate activity-dependent gene transcription. Mutations of proteins involved in activity-dependent gene expression are associated with some forms of ASD, including L-VSCC in Timothy syndrome, RSK2 in Coffin–Lowry syndrome, CBP in Rubinstein-Taybi syndrome, Ube3A in Angelman syndrome and MECP2 in Rett syndrome. Ac, acetylation; CaRF, calcium response factor; Me, methylation; P, phosphorylation.

the number of genes for which mutation may contribute to ASD. As *de novo* mutations within gene-coding regions occur with reasonable frequency and do not necessarily result in a deleterious phenotype, additional evidence is required before each of these mutations can unequivocally be said to contribute to ASD. In the future, the key to understanding how mutations in these proteins lead to ASD will be to identify additional individuals with ASD, who carry mutations in these genes, and to obtain a more thorough understanding of the molecular functions of each of these proteins in neurons and how their mutation leads to defects in synaptic function.

In addition to these rare mutations, several widely studied syndromic disorders, such as fragile X syndrome, with penetrant ASD features have provided useful insight into the pathophysiology of ASD. These syndromic disorders are caused by single gene mutations and have phenotypes that have allowed the disorder to be readily diagnosed; thus, the mechanistic understanding of these syndromes is more advanced than for most forms of ASD. For example, fragile X syndrome — characterized by macroorchidism, connective-tissue disease, motor abnormalities, language deficits and impaired social interactions — is the leading cause of inherited intellectual disability and ASD[23]. Fragile X syndrome results from a trinucleotide expansion (CGG) in *FMR1*, which encodes fragile X mental retardation protein (FMRP). The CGG expansion occurs within the 5′ untranslated region of the FMRP mRNA, resulting in decreased translation of the FMRP mRNA and a decrease in the amount of FMRP expressed in neurons. FMRP has recently been reported to modulate the rate of translational elongation of mRNAs that encode synaptic proteins[33], suggesting that the decrease in FMRP expression that occurs in fragile X syndrome may result in the dysregulation of synaptic protein expression owing to an impairment in translational control[23]. Thus, as described later, studying the genes that are mutated in these widely studied syndromes has the potential to provide insight into the molecular and cellular basis of ASD.

Several genetic loci (*UBE3B*, *CLTCL1*, *NCKAP5L* and *ZNF18*) that are associated with ASD include genes for which expression is induced by neuronal activity[34,35]. In addition, knockdown of several transcription factors that mediate activity-regulated gene transcription, including MEF2 and NPAS4, has been shown to lead to decreased expression of multiple genes that are linked to ASD (*c3orf58*, *SLC9A9* and *PCDH10*)[34]. Only about 3% of the neuronal transcriptome is altered when the membrane is depolarized, but a higher percentage of the genes that are implicated in ASD have been shown to be regulated by neuronal activity; it has been proposed, therefore, that ASD may result, at least in part, from the disruption of the activity-dependent signalling networks that control gene transcription[34,35]. Thus, it will be important to determine whether the expression of RNAs that are encoded by other genetic loci associated with ASD are regulated by neuronal activity. A crucial consideration is that genetic loci associated with ASD may not only be found to disrupt gene function through mutations within gene-coding sequences, but may also be found to disrupt the function of the regulatory regions that control these genes. Of note, therefore, is a recent study that identified thousands of neuronal activity-regulated enhancers of transcription that could potentially be sites of human variation that lead to ASD[36].

Synaptic proteins implicated in ASD

Mutations in multiple synaptic cell-adhesion molecules and components of the PSD are associated with ASD. The fact that these proteins have key roles in regulating synaptic responses to experience, and are themselves modulated by neuronal activity supports the hypothesis that ASD may result from the disruption of the normal process of activity-dependent synapse development (Fig. 2).

Neuroligins and neurexins

Rare mutations in multiple members of the neuroligin and neurexin families of synaptic adhesion molecules have repeatedly been found to be associated with ASD[5]. Neurexins are presynaptic membrane proteins that are encoded by mRNAs that have undergone extensive alternative splicing. Neuroligins are postsynaptic membrane proteins that function as neurexin receptors and, working together with neurexins, modulate the formation and function of synapses[16,37]. Particular neuroligin family members are localized specifically to inhibitory or excitatory synapses, suggesting that individual neuroligins may play a part in specifying the formation or function of excitatory and inhibitory synapses. In addition to binding neuroligins, neurexins bind other postsynaptic transmembrane proteins, including leucine-rich repeat (LRR)-containing transmembrane proteins, adding to the number of different potential *trans*-synaptic connections[16]. There is increasing evidence for the involvement of neurexins, neuroligins and LRR-containing transmembrane proteins in activity-dependent programs that modulate synapse function.

Several of the neuroligin mutations found in individuals with ASD have been modelled in mice, and these mice have been shown to have alterations in excitatory or inhibitory neurotransmission. For example, knock-in mice that harbour the neuroligin-3 ASD missense mutation *Nlgn3*(R451C) display impaired social interactions, recapitulating a key feature of ASD[38]. These knock-in mice also have enhanced inhibitory neurotransmission with no alterations in excitatory neurotransmission, resulting in a defect in excitatory–inhibitory balance in the brain[38]. Knock-in mice with another ASD-associated missense mutation, *Nlgn3*(R704C) displayed a decrease in AMPA-receptor-mediated synaptic transmission in the hippocampus, but no alteration in NMDA-receptor or GABA (γ-aminobutyric acid)-receptor-mediated neurotransmission[39]. In addition to the neuroligins, mutations in *NRXN1* are also associated with ASD in humans. *Nrxn1* knockout mice have defects in miniature excitatory postsynaptic current (EPSC) frequency and evoked postsynaptic potentials, but show no change in inhibitory neurotransmission in the hippocampus[40]. These studies indicate that particular ASD-associated mutations in neurexins or neuroligins disrupt excitatory or inhibitory neurotransmission in the brain in specific ways and suggest that modelling ASD-associated mutations in mice will be important for elucidating how each mutation affects synaptic function and gives rise to ASD.

Notably, neuroligins have been found to regulate the activity-dependent maturation and modulation of synapses. In cultured neurons, overexpression of *Nlgn1* in rodents increases the number of functional excitatory synapses, an effect that is dependent on NMDA receptors, as well as activity-regulated CaMKII, which functions downstream of NMDA receptor activation to regulate synaptic plasticity[41]. In contrast to overexpression of wild-type *Nlgn1*, overexpression of NLGN1 containing an ASD-associated missense mutation leads to a decrease in the number and strength of excitatory synapses[41]. This suggests that activity-dependent functions of NLGN1 may be disrupted in ASD. In another study[42], knockdown of *Nlgn1*, delivered by a viral vector, in the amygdala of adult rats led to the selective reduction of NMDA-receptor-mediated currents and to impairments in LTP and in the storage of associative fear memory. This suggests a role for neuroligins in experience-regulated learning and memory, which may be altered by ASD-associated mutations in *Nlgn1*.

Similarly, using sensitization of the gill-withdrawal reflex in the sea hare *Aplysia californica* as a simple model of learned fear and long-term memory showed that depletion of the *A. californica* homologue of neuroligin in the postsynaptic motor neuron, or neurexin in the presynaptic sensory neuron, impaired initiation, stabilization and persistence of long-term facilitation of the sensory-to-motor neuron synaptic connection. Relative to wild-type neuroligin, overexpression of neuroligin with an ASD-associated mutation blocked intermediate and long-term facilitation at the sensory-to-motor neuron synapse[43]. Taken together, these findings suggest that neuroligins and neurexins modulate activity-dependent synaptic remodelling and long-term memory, and that ASD-associated mutations in neuroligins or neurexins may disrupt these processes.

The interaction between neurexins and neuroligins is regulated by neuronal activity, providing further evidence that mutations that cause ASD might be disrupting activity-dependent processes that control synaptic function. The dynamic nature of the neuroligin–neurexin interaction was studied using the enzyme biotin ligase (BirA) from the bacterium *Escherichia coli*, which catalyses the biotinylation of an acceptor peptide only when BirA and the acceptor peptide are in close proximity[44]. When neurexin-1 fused to BirA and neuroligin fused to an acceptor peptide are expressed in neurons, biotinylation of neuroligin–acceptor protein occurs when neurexin-1–BirA and neuroligin–acceptor protein interact across the synapse. Using this approach, neuronal activity was found to acutely promote an increase in the interaction of neurexins and neuroligins across the synapse in an NMDA-receptor-dependent manner[44]. Neuronal activity induced both the insertion of neurexin and neuroligin into the synaptic membrane and suppressed internalization of neurexin and neuroligin from the synaptic membrane, leading to an overall increase in the *trans*-synaptic neurexin–neuroligin interaction. Disruption of this activity-regulated enhancement impaired the recruitment of AMPA receptors to the synaptic membrane, thereby altering synaptic function[44]. It will be important to determine whether ASD-associated mutations in neurexins and neuroligins affect these activity-dependent changes in their *trans*-synaptic interaction.

Neuronal activity also regulates alternative splicing of neurexin mRNA and, thus, alters *trans*-synaptic signalling and presynaptic differentiation[45]. It will be interesting to determine if ASD-associated mutations disrupt activity-dependent alternative splicing of neurexin and whether disruption of this process contributes to the development of ASD.

SHANKs

Shank proteins function as scaffolds in the PSD of excitatory synapses that regulate the organization of postsynaptic signalling complexes, as well as the morphology and function of synapses[17–19]. Rare mutations associated with ASD have been found in *SHANK2* and *SHANK3* (ref. 5), and recently in *SHANK1* (ref. 46). In particular, Phelan–McDermid syndrome is an ASD that results from a microdeletion that includes heterozygous loss of *SHANK3*. Knockout mice with deletions in members of the Shank family of genes exhibit behaviours similar to those observed in ASD. For example, *Shank3* knockout mice have deficits in social interactions and engage in repetitive behaviours, such as excessive grooming, that lead to self injury[47]. In addition, mice lacking *Shank2* have abnormalities in motor behaviour, vocalization and socialization — phenotypes typically observed in ASD[48,49].

These *Shank2* and *Shank3* knockout mice display alterations in excitatory neurotransmission that correlate with ASD behavioural phenotypes. The *Shank3* knockout mice have deficits in striatal glutamatergic synapse structure and function, including reduced cortico-striatal synaptic transmission[47], whereas *Shank2* knockout mouse models show defects in NMDA-receptor-dependent excitatory neurotransmission and synaptic plasticity in the hippocampus[48,49]. In one study of *Shank2* knockout mice, by using either a partial agonist of NMDA receptor function (D-cycloserine) or a positive allosteric modulator of mGluR5, it was possible to normalize NMDA receptor function and decrease autistic behaviours, suggesting that impairment of NMDA receptor functioning may be a key mechanism through which mutations in genes that encode the SHANK family of proteins lead to ASD[49]. Taken together, these studies have revealed how the SHANK family of proteins modulate excitatory synapse function and begin to explain how mutations in the genes that encode them might lead to ASD.

One possibility is that SHANK mutations that cause ASD disrupt activity-dependent regulation of SHANK function. Neuronal activity has been shown to affect the molecular composition of the postsynaptic compartment by inducing the ubiquitination and proteasome-mediated degradation of a subset of PSD proteins, including those of the Shank family, thus leading to changes in synaptic signalling[50,51]. The ubiquitin ligase TRIM3 mediates the activity-dependent degradation of SHANK1 and another PSD-scaffolding protein GKAP, by a ubiquitin–proteasome-dependent mechanism[52]. In mice, the retrieval of fear memory was found to lead to the polyubiquitination and degradation of SHANK proteins and to correlate with a requirement for protein degradation in memory extinction[53]. These findings suggest that the activity-dependent degradation of synaptic proteins, including those of the SHANK family, regulates the reorganization of synapses and memories. Taken together, these studies begin to explain how SHANK proteins control synaptic function and plasticity, and how their mutation might contribute to ASD phenotypes.

Activity-dependent regulation of mRNA translation

Long-lasting forms of synaptic plasticity that are elicited by glutamate binding to NMDA receptors or group 1 mGluRs, typically require protein synthesis mediated by mRNAs and ribosomes that are localized near synapses. Long-lasting NMDA-receptor-dependent LTP and LTD have a delayed requirement for protein synthesis, whereas mGluR-dependent LTD requires protein synthesis that occurs within minutes. Several genes that are mutated in ASD — *FMR1*, *TSC1*, *TSC2* and *PTEN* — have key roles in protein-synthesis-dependent plasticity of synapses[13,23] (Fig. 2).

Fragile X syndrome

Fragile X syndrome is caused by mutations in *FMR1*, resulting in decreased expression of the RNA-binding protein FMRP. This protein negatively regulates the translation of specific mRNAs at the synapse[23,54]. FMRP has been suggested to function by blocking the initiation of mRNA translation, or stalling movement of the ribosome along the mRNA, leading to decreased protein synthesis[33]; however, the precise molecular mechanisms by which FMRP regulates translation remain unclear[23]. *Fmr1* knockout mice exhibit abnormally high levels of protein synthesis and enhanced LTD in response to mGluR stimulation, suggesting that under normal conditions FMRP suppresses mGluR-dependent protein synthesis and LTD[55]. Remarkably, the excessive protein synthesis observed in *Fmr1* knockout mice is reversed by antagonists of mGluR5. In addition, both genetic and pharmacological inhibition of mGluR5 reverses synaptic and behavioural impairments in these mice. These findings are the basis of human clinical trials that seek to test the efficacy of mGluR5 antagonists for treating fragile X syndrome[23,56–58].

Arc is one of the best-studied mRNA targets of FMRP. The protein encoded by this activity-regulated gene promotes the internalization of AMPA receptors at excitatory synapses. Translation of *Arc* mRNA is enhanced during, and crucial for, mGluR-dependent LTD[59,60]. In *Fmr1* knockout mice, the dynamic translation of *Arc* mRNA in response to stimulation is impaired[59,61], and the basal level of the ARC protein in dendrites is elevated[61]. This abnormally high level of *Arc* expression in *Fmr1* knockout mice may be responsible for the enhanced mGluR–LTD observed in these animals and might explain, in part, why their LTD is independent of new mRNA translation[62].

During early postnatal brain development, neuronal activity promotes changes in glutamate receptor subtype that are required for proper maturation of excitatory thalamocortical synapses in the somatosensory cortex. In *Fmr1* knockout mice, this activity-dependent maturation of the excitatory thalamocortical synapses is dysregulated, resulting in a persistence of silent NMDA-receptor-only synapses at times during brain development when these synapses would normally express both NMDA and AMPA receptors[63]. The absence of AMPA receptors at these synapses may be due, in part, to the dysregulation of *Arc* mRNA translation that occurs in *Fmr1*-knockout mice.

FMRP has also been suggested to be crucial for activity-dependent synapse elimination, a key process during postnatal brain development[64] that may be defective in fragile X syndrome. The transcription factor MEF2 controls a program of gene expression that has a role in developmental synapse elimination[26,29,65,66]. Notably, activation of MEF2 leads to a decrease in the number of excitatory synapses in wild-type neurons, but not in *Fmr1* knockout neurons, suggesting that the effects of MEF2 on synapse number may be FMRP dependent[64]. ARC may be the link between MEF2 and FMRP. MEF2 seems to restrict excitatory synapse number, at least in part, by activating *Arc* transcription. Once *Arc* transcription is induced, translation of the *Arc* mRNA is kept in check by FMRP. However, in the absence of FMRP, the level of ARC protein is abnormally high, so further activation of *Arc* transcription by MEF2 may not have much of an effect on synapse number.

However, ARC is almost certainly not the only target of FMRP that mediates the effects of FMRP on LTD, excitatory synapse number and ASD. Many other FMRP targets have been shown to be involved in synaptic plasticity, and there is significant overlap between FMRP targets and ASD candidate genes, including *NLGN3*, *NRXN1*, *SHANK3*, *PTEN*, *TSC1* and *NF1* (refs 23, 33). This suggests that these ASD candidate genes and FMRP may function in overlapping pathways. Taken together, these studies implicate FMRP as a key regulator of activity-dependent mRNA translation and synaptic plasticity and suggest that dysregulation of this pathway may contribute significantly to ASD phenotypes.

Tuberous sclerosis complex

Tuberous sclerosis complex is the other main syndromic ASD that seems to be a consequence of a dysregulation of activity-modulated mRNA translation. Tuberous sclerosis complex is characterized by benign tumours in multiple organs, as well as a high penetrance of ASD. This disorder is caused by mutations in TSC1 or TSC2, which encode tumour suppressor proteins that form the complex TSC1–TSC2 and regulate protein synthesis in a wide range of mammalian cell types, including neurons[67]. The binding of growth factors to receptor tyrosine kinases leads to phosphorylation of the TSC protein complex and subsequent activation of mTOR, a positive regulator of protein synthesis[67].

One major activator of TSC1–TSC2 signalling in neurons is brain-derived neurotrophic factor (BDNF), a secreted protein that binds to the receptor tyrosine kinase TrkB on the surface of neurons and thereby mediates many aspects of brain development and function. Neuronal activity potently induces *Bdnf* transcription in excitatory neurons. Newly synthesized BDNF is then secreted from excitatory neurons and functions near the site of release to regulate synaptic development and plasticity, including the development of inhibitory synapses[30]. Several studies suggest that the effect of BDNF on these processes is mediated by the local control of protein translation[68]. Specifically, BDNF binding to TrkB activates phosphatidylinositol-3-OH kinase (PI(3)K), which phosphorylates phosphatidylinositol-4,5-bisphosphate (PtdIns(4,5)P_2) at the 3′ position to generate phosphatidylinositol-3,4,5-trisphosphate (PtdIns(3,4,5)P_3). This then binds and activates phosphoinositide-dependent protein kinase (PDK1), which in turn contributes to activation of the serine/threonine protein kinase AKT proteins. Activated AKT phosphorylates TSC2, reducing its GTPase-activating protein (GAP) activity for the Ras family GTPase, Rheb. Thus, phosphorylation of TSC2 leads to Rheb activation, which in turn triggers activation of the protein kinase mTOR, contributing to the activation of protein synthesis[67] (Fig. 4). Loss-of-function mutations in *TSC1* or *TSC2* that occur in tuberous sclerosis complex, result in deregulated mTOR signalling and abnormal neuronal protein synthesis, suggesting that tuberous sclerosis complex may be a result of the dysregulation of these processes[13,69]. Although it remains to be demonstrated that the disruption of BDNF-mediated protein synthesis has a key role in TSC1–TSC2-dependent ASD, the observations that BDNF has a widespread role in nervous system development, that *Bdnf* transcription is robustly induced by neuronal activity and that BDNF regulates protein translation through inhibition of TSC1–TSC2 suggest a role for BDNF in tuberous sclerosis complex. Signalling by TSC1–TSC2 downstream of other neurotrophic factors is also likely to contribute to the disorder.

Further evidence that a defect in neurotrophic-factor-regulated protein synthesis contributes to ASD is provided by the finding that mutations in *PTEN*, a negative regulator of PI(3)K–AKT signalling upstream of TSC1–TSC2, also result in syndromic ASD. The lipid phosphatase PTEN catalyses the dephosphorylation of PtdIns(3,4,5)P$_3$ at the 3′ position to generate PtdIns(4,5)P$_2$, thus antagonizing the activity of PI(3)K[70]. Mutations that inactivate PTEN lead to excessive PI(3)K–AKT and mTOR signalling and in this way may lead to abnormal protein synthesis, and ultimately synaptic dysfunction and ASD. Thus, characterization of the role of PTEN in neurotrophic-factor-regulated protein synthesis may help to uncover the mechanisms that contribute to ASD.

The study of mice that carry mutations in *Pten*, *Tsc1* or *Tsc2* has the potential to provide insight into how defects in PTEN–PI(3)K–AKT–TSC signalling lead to ASD. Loss of function of TSC1 and TSC2 leads to deregulated activation of mTOR[13,69]. As TSC1 and TSC2 are negative regulators of protein synthesis, it was reasonable to hypothesize that — similar to loss of FMRP — loss of TSC1 or TSC2 would lead to increased protein synthesis and enhanced mGluR–LTD. Interestingly, the opposite has been found in recent studies[71,72]. Mice with a heterozygous loss-of-function mutation in *Tsc2* (*Tsc2*$^{+/-}$) have deficient mGluR–LTD in the hippocampus, but no defect in basal synaptic transmission. Compared with wild-type mice, *Tsc2*$^{+/-}$ mice have decreased translation of *Arc* mRNA[71]. Similarly, in another study, postnatal deletion of *Tsc1* in hippocampal CA1 mouse neurons by viral delivery of Cre recombinase led to a loss of mGluR–LTD[72]. These findings of deficient mGluR–LTD in mice with deletions in *Tsc1* or *Tsc2* are in contrast to those in *Fmr1* knockout mice, which have increased mGluR–LTD and excessive protein synthesis[57].

Inhibition of mTOR in *Tsc2*$^{+/-}$ mice restored mGluR–LTD and normalized protein synthesis rates to those of wild-type mice, suggesting that excessive mTOR activity leads to decreased protein synthesis and LTD deficiency in *Tsc2*$^{+/-}$ mice[71]. Whereas inhibition of mGluR5 in *Fmr1* knockout mice corrects excessive protein synthesis and mGluR5-dependent LTD, augmentation of mGluR5 signalling by a positive allosteric modulator restored protein synthesis and mGluR-dependent LTD in *Tsc2*$^{+/-}$ mice[71]. In a fear-conditioning behavioural model, *Tsc2*$^{+/-}$ mice have a defect in distinguishing familiar and new contexts, a defect that can also be corrected by augmentation of mGluR5 signalling[71]. These findings suggest that, for proper synaptic function in response to neuronal activity, protein synthesis at synapses is tightly regulated, and *Tsc1* and *Tsc2* as well as *Fmr1* mutations disrupt this balance in opposite directions, leading to defects in synaptic plasticity that may contribute to ASD phenotypes.

Despite recent progress, major gaps in knowledge remain about how neuronal activity dynamically regulates protein synthesis to control synaptic function and how disruption of TSC1–TSC2 or FMRP alters these processes to lead to ASD. Indeed, how loss of function of FMRP or TSC1–TSC2 in fragile X syndrome and tuberous sclerosis complex leads to differential dysregulation of protein synthesis and synaptic plasticity is not yet understood. One possibility is that enhanced mTOR activity in tuberous sclerosis complex results in increased FMRP phosphorylation, which in turn functions to suppress the translation of specific mRNAs. Indeed, stimulus-regulated phosphorylation or dephosphorylation of FMRP alters its ability to suppress mRNA translation[61,73–75]. Alternatively, TSC1–TSC2 and FMRP may regulate the translation of different pools of mRNAs, so that abnormal translation of the two pools of mRNAs has different outcomes. Improved understanding of the activity-dependent programs that control protein synthesis and synaptic plasticity will probably help to clarify the molecular mechanisms that underlie these two forms of ASD.

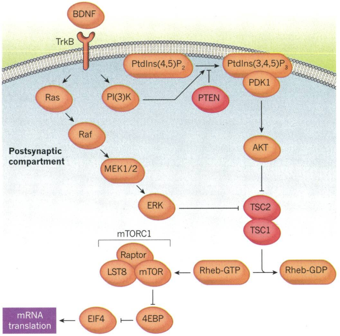


Figure 4 | BDNF regulates mRNA translation. Neuronal activity potently induces expression of the neurotrophic factor BDNF. BDNF binds the receptor TrkB, a receptor tyrosine kinase, triggering several signalling cascades, including the PI(3)K–AKT and Ras–MAPK pathways. PI(3)K is a lipid phosphatase, converting PtdIns(4,5)P$_2$ to PtdIns(3,4,5)P$_3$, which can be reversed with the phosphatase PTEN. PDK1 together with PtdIns(3,4,5)P$_3$ activates AKT, which inhibits TSC1–TSC2. TSC1–TSC2 serves as a GTPase activating protein for the GTPase Rheb, promoting the conversion of active GTP-bound Rheb to the inactive GDP-bound form. Rheb-GTP activates mTOR to induce mRNA translation. PTEN and TSC1–TSC2 are associated with ASD, as are mutations in Ras, Raf and MEK1–MEK2.

Activity-dependent gene transcription

In addition to its effects on protein synthesis at synapses, neuronal activity regulates programs of gene expression in the nucleus that modulate synaptic development and plasticity. Mutations associated with ASD disrupt multiple features and components of the signalling network that controls activity-dependent gene transcription. These include effects on the route and dynamics of calcium entry into neurons, effects on transcriptional regulators that control gene expression in response to neuronal activity, and an impact on effector genes that encode proteins that regulate synaptic development and function (Fig. 3).

Timothy syndrome

The route and dynamics of calcium entry into the neuron determine the program of activity-dependent gene expression that is activated[24,25]. In excitatory neurons, activity-dependent gene transcription is mediated, at least in part, by calcium entry through L-type calcium channels[25]. The L-type calcium channels are voltage-gated and thus open upon membrane depolarization. Mutations in *CACNA1C*, which encodes the α1 pore-forming subunit of the L-type calcium channel $Ca_v1.2$, have been associated with multiple neuropsychiatric disorders, including ASD[76]. In addition, Timothy syndrome, a disorder characterized by cardiac arrhythmia, congenital heart disease, immune deficiency, intermittent hypoglycaemia, global development delay and ASD, is caused by a missense mutation in *CACNA1C*[31]. Studies in mice indicate that this missense mutation leads to increased Ca^{2+} currents and loss of voltage-dependent channel inactivation[31,77]. To determine the effect of the Timothy syndrome mutation on human neurons, induced pluripotent stem cells were derived from fibroblasts of individuals with Timothy syndrome and differentiated into neurons[78]. Compared with wild-type neurons, the Timothy syndrome neurons had impaired Ca^{2+} channel inactivation and a significant increase in sustained intracellular Ca^{2+} elevation after membrane depolarization. Activity-dependent gene expression was also altered in the Timothy syndrome neurons, compared with control neurons, with more than 200 genes displaying altered expression following membrane depolarization[78]. These findings raise the possibility that the ASD phenotypes associated with Timothy syndrome may be the result of the deregulation of activity-dependent gene transcription.

Rett syndrome

Mutations in *MECP2* lead to Rett syndrome, which is a form of ASD characterized by impaired language development, loss of social engagement, stereotyped hand movements, seizures and motor-system disabilities. MECP2 binds methylated cytosines within DNA and seems to function mainly as a transcriptional repressor. The level of MECP2 expression in post-mitotic neurons is about tenfold higher than in other cell types, suggesting that MECP2 may have a unique function in neurons. Notably, the level of MECP2 expression in mature neurons is similar to that of core histones, and MECP2 is bound broadly across the genome. These findings suggest that MECP2 may have a histone-like role in regulating chromatin structure and gene expression[79,80]. Neuronal activity induces the phosphorylation of MECP2 at Ser 421 (refs 81–83) broadly across the genome, raising the possibility that activity-dependent phosphorylation of MECP2 mediates a genome-wide chromatin response to neuronal activity[80]. The fact that Rett syndrome manifests itself in the first year of life — a time when experience is shaping synapse development — together with the observation that MECP2 is a chromatin regulator that is modified in response to neuronal activity, suggests that Rett syndrome may be due to a defect in the process of activity-dependent chromatin regulation or gene transcription that controls synaptic development, plasticity and behaviour.

To test the role of activity-dependent phosphorylation of MECP2 in synapse development and behaviour, knock-in mice were generated in which Ser 421 was converted to an alanine (Ser 421Ala), preventing this residue from being phosphorylated[80,84]. The Ser 421Ala knock-in mice and wild-type mice express equivalent levels of MECP2, but the Ser 421Ala knock-in mice display increased dendritic complexity and increased inhibitory synaptic strength in the cortex[80]. Behaviourally, the Ser 421Ala knock-in mice had deficits in their response to new objects or mice[80]. These findings indicate that activity-dependent phosphorylation of MECP2 regulates synapse development and function and behavioural responses to environmental stimuli. However, extensive analysis of activity-dependent gene transcription in wild-type and MECP2 Ser 421Ala knock-in neurons failed to detect any effect of the Ser 421Ala mutation. Thus, it remains to be determined whether activity-dependent phosphorylation of MECP2 affects gene transcription and chromatin structure and, if so, how the disruption of these processes contributes to Rett syndrome.

MECP2 has recently been shown to be required for experience-dependent synaptic remodelling of the retinogeniculate synapse[85]. In *Mecp2* knockout mice, the initial development and refinement of the retinogeniculate circuit proceeds without impairment, during a phase of development that does not depend on visual experience. However, during a late stage of development when visual experience is required for modulation and maintenance of the retinogeniculate circuit, the synapses that form on the neurons within this circuit are abnormal in *Mecp2* knockout mice, resulting in impaired synaptic plasticity on visual deprivation[85]. This raises the possibility that during the late phase of development of the retinogeniculate circuit, MECP2 may regulate an experience-dependent gene expression program that controls synaptic remodelling, a program that could potentially be controlled by activity-dependent phosphorylation of MECP2.

There is also evidence that loss of MECP2 changes the balance between inhibition and excitation in the brain, altering the neuronal response to sensory experience. In *Mecp2* knockout mice, whole-cell patch clamp recordings in the somatosensory cortex revealed decreased excitatory synaptic drive and increased inhibitory drive onto L5 pyramidal cells[86]. Hippocampal neurons from *Mecp2* knockout mice, both grown in culture and *in vivo*, have reduced glutamatergic synapse number and density with reduced excitatory synapse strength[87]. In addition, mice expressing a truncated *Mecp2* variant, lacking the carboxy-terminal region, have multiple defects in learning and memory, as well as synaptic plasticity[88]. Together with the molecular studies of MECP2 function, these electrophysiological experiments suggest that MECP2 is a key regulator of chromatin architecture and gene expression that modulates synaptic development, function and plasticity. It is likely that the dysregulation of at least some of MECP2's functions leads to Rett syndrome and possibly other forms of ASD.

Angelman syndrome

Angelman syndrome is characterized by motor dysfunction, speech impairment, seizures and a high prevalence of autism, and is caused by deletions of the maternally inherited allele of chromosomal region 15q11–q13. The loss of *UBE3A* function in this chromosomal region is by itself sufficient to cause Angelman syndrome. Notably, CNVs that lead to duplications and triplications of the same chromosomal region are associated with ASD. *Ube3a* expression is upregulated by neuronal activity; it has therefore been suggested that a loss of experience-dependent induction of the *UBE3A* mRNA during postnatal development may contribute to Angelman syndrome, and increased activity-dependent *UBE3A* expression may underlie some forms of ASD[89]. Support for the idea that fine-tuning the level of *UBE3A* expression may be important for normal cognitive function comes from the observation that the paternally inherited allele of *UBE3A* is imprinted, and is specifically silenced in the brain. Thus, the level of the gene's expression in the nervous system is kept in check through imprinting and is further regulated by neuronal activity.

Mice that carry a deletion in the maternal copy of *Ube3a* have impairments in learning and in the LTP form of synaptic plasticity[90]. Experience-dependent maturation of excitatory circuits in the visual cortex is substantially impaired in these same mice. In addition, sensory experience-regulated neocortical plasticity is defective in *Ube3a*-mutant mice[91]. Specifically, inhibitory drive onto neocortical pyramidal neurons is impaired owing to defective presynaptic vesicle cycling in interneuron populations[92]. Excitatory drive onto neocortical pyramidal neurons is also diminished, but to a less severe extent[92]. These studies suggest that UBE3A regulates excitatory–inhibitory balance during brain development and that when the level of expression is perturbed ASD may occur.

UBE3A is a member of a family of E3 ubiquitin ligases, which catalyse the ubiquitination of substrates, often leading to the degradation of the ubiquitinated substrate by the proteasome. Notably, one of the best-studied targets of UBE3A is the neuronal activity-regulated protein ARC[89], suggesting a possible convergence point for UBE3A, FMRP and TSC mutations in the aetiology of ASD. Studies using mice that carry a deletion in maternally inherited *Ube3a* revealed that neurons from these mice express increased levels of ARC protein. In addition, UBE3A was found to mediate the ubiquitination and degradation of ARC[89]. The increased level of ARC expression in the absence of UBE3A results in decreased AMPA receptor cell-surface expression and a decrease in AMPA-receptor-mediated currents[89]. Thus, the dysregulation of ARC levels and AMPA-receptor expression at synapses may contribute to the ASD phenotypes observed in Angelman syndrome and possibly forms of ASD that involve CNVs on chromosome 15.

The studies described here on Angelman syndrome, Timothy syndrome and Rett syndrome suggest that the dysregulation of molecular pathways that control activity-dependent gene expression — including those that control the influx of calcium into neurons (CACNA1C in Timothy syndrome), the regulation of chromatin and transcription (MECP2 in Rett syndrome) and the expression of effector genes that act on synapses (UBE3A in Angelman syndrome) — may be crucial to ASD pathogenesis. Thus, it seems likely that further characterization of the signalling network that controls activity-dependent gene expression in neurons will reveal additional genes that are mutated in ASD.

Future directions

The studies reviewed here suggest that dysregulation of activity-dependent signalling networks that control synapse development, function and plasticity may contribute significantly to ASD pathogenesis. A variety of studies now provide evidence that mutations that lead to ASD affect how sensory experience modifies synapses, through local changes in protein phosphorylation at the synapse, the control of mRNA translation or the regulation of gene transcription. However, further experimentation will be required to provide definitive evidence that disruption of activity-dependent programs in neurons is an underlying cause of ASD. In addition, many questions remain to be addressed regarding the molecular mechanisms by which mutations associated with ASD disrupt the activity-dependent pathways that control synapse development, maturation and function.

It will be important to determine whether perturbation of activity-dependent gene regulation contributes to ASD pathology. This could be addressed by evaluation of the effects of subtle mutations, such as knock-in mutations in promoters or enhancers, that specifically disrupt activity-dependent gene transcription. Similarly, it will be useful to determine whether dysregulation of activity-dependent splicing or translation of mRNAs that are implicated in ASD have a role in the development of the disorder, and whether disruption of the activity-dependent phosphorylation of specific synaptic proteins leads to ASD phenotypes. For instance, analysis of knock-in mice that carry subtle mutations that disrupt activity-dependent phosphorylation of specific proteins may prove useful in addressing these issues.

To further advance our understanding of ASD, it will also be crucial to continue to investigate the function of ASD candidate genes, particularly those that are part of the activity-dependent signalling network that controls synapse development and function. In this regard, it is important to keep in mind that although not all ASD candidate genes are part of the activity-dependent signalling network, careful examination may reveal additional ASD-associated genes that regulate various aspects of activity-dependent signalling. For example, ASD-associated mutations may disrupt activity-dependent protein synthesis or degradation at or near the synapse. In addition, ASD-associated mutations could disrupt activity-dependent protein localization within the neuron, such as recruitment to the synapse or protein clustering at the synaptic cleft, or neuronal activity-dependent protein phosphorylation events that control channel conductance properties or protein–protein interactions. ASD-associated mutations could also disrupt routes and dynamics of calcium influx into neurons or affect signalling complexes that transmit calcium-dependent signals to control gene transcription. As new ASD-associated loci are identified, it will be important to determine which loci encode proteins that are regulated by neuronal activity, how neuronal activity affects the function of the protein and whether the effect on protein function is relevant to the ASD phenotype. One signalling pathway of significant interest in this regard is Ras–MAPK. Mutations in multiple components of this pathway have recently been found to be associated with ASD, including rare mutations in *HRAS*, *MAP2K2* and *SYNGAP1*, and with syndromes with ASD features (mutations in *NF1*, *MAP2K1* and *MAP2K2*)[93,94]. As the Ras–MAPK signalling pathway is crucial for activity-regulated protein synthesis, activity-dependent gene transcription and mGluR-dependent synaptic plasticity[23,25,95], mutations in this pathway that are associated with ASD may be disrupting activity-dependent synaptic development and plasticity by altering activity-regulated mRNA translation or gene transcription.

Understanding the molecular basis of ASD will hopefully allow rational development of therapies for these disorders. Several mouse models of ASD, including those of Rett syndrome[96,97], Angelman syndrome[98], tuberous sclerosis complex[71], and fragile X syndrome[23], and mouse models of rare mutations in SHANK proteins[49], have provided convincing evidence that ASD phenotypes can be improved by genetic or pharmacological interventions, providing hope that it will be possible to develop therapeutics for ASD that target the underlying pathophysiology and provide significant clinical benefit. Defining how each type of ASD-associated mutation disrupts synaptic function may be crucial for the development of effective treatments that target the specific molecular defect that occurs in each disorder.

A more general therapeutic approach might be to target a cellular phenotype, such as the imbalance in excitatory–inhibitory neurotransmission that seems to be common to many forms of ASD. The mGluR5 antagonists that are being evaluated for efficacy in fragile X syndrome are an example of therapeutics that have been developed to treat a specific disorder, but that might have more general applicability. Pharmacological screens for small molecules that correct deficits in activity-dependent signalling or excitatory–inhibitory balance may also prove useful. For instance, small molecules that target regulators of phosphorylation-mediated or ubiquitin-mediated signalling in neurons may be able to counter the dysregulation of activity-dependent signalling that occurs in certain forms of ASD. Thus, a better understanding of the role of neuronal activity-dependent signalling in the aetiology of ASD may be useful for both advancing our understanding of the molecular basis of ASD and developing new therapeutics for this cognitive disorder. ■

Received 14 June; accepted 8 November 2012.

1. Baird, G. *et al.* Prevalence of disorders of the autism spectrum in a population cohort of children in South Thames: the Special Needs and Autism Project (SNAP). *Lancet* **368,** 210–215 (2006).
2. Zoghbi, H. Y. & Bear, M. F. Synaptic dysfunction in neurodevelopmental disorders associated with autism and intellectual disabilities. *Cold Spring Harb. Perspect. Biol.* **4,** a009886 (2012).
3. Flavell, S. W. & Greenberg, M. E. Signaling mechanisms linking neuronal activity to gene expression and plasticity of the nervous system. *Annu. Rev. Neurosci.* **31,** 563–590 (2008).
4. Abrahams, B. S. & Geschwind, D. H. Advances in autism genetics: on the threshold of a new neurobiology. *Nature Rev. Genet.* **9,** 341–355 (2008).
5. State, M. W. & Levitt, P. The conundrums of understanding genetic risks for autism spectrum disorders. *Nature Neurosci.* **14,** 1499–1506 (2011).
6. Malhotra, D. & Sebat, J. CNVs: harbingers of a rare variant revolution in psychiatric genetics. *Cell* **148,** 1223–1241 (2012).
7. Sanders, S. J. *et al. De novo* mutations revealed by whole-exome sequencing are strongly associated with autism. *Nature* **485,** 237–241 (2012).
8. Neale, B. M. *et al.* Patterns and rates of exonic *de novo* mutations in autism spectrum disorders. *Nature* **485,** 242–245 (2012).
9. O'Roak, B. J. *et al.* Exome sequencing in sporadic autism spectrum disorders identifies severe *de novo* mutations. *Nature Genet.* **43,** 585–589 (2011).
10. Kong, A. *et al.* Rate of *de novo* mutations and the importance of father's age to disease risk. *Nature* **488,** 471–475 (2012).
    References 7–10 demonstrate the ability of new sequencing technologies to significantly advance our understanding of the genetic basis of ASD.
11. Kandel, E. R. The molecular biology of memory storage: a dialogue between genes and synapses. *Science* **294,** 1030–1038 (2001).
12. Greer, P. L. & Greenberg, M. E. From synapse to nucleus: calcium-dependent gene transcription in the control of synapse development and function. *Neuron* **59,** 846–860 (2008).

13. Kelleher, R. J. III & Bear, M. F. The autistic neuron: troubled translation? *Cell* **135,** 401–406 (2008).
14. Wiesel, T. N. Postnatal development of the visual cortex and the influence of environment. *Nature* **299,** 583–591 (1982).
15. Cohen, S. & Greenberg, M. E. Communication between the synapse and the nucleus in neuronal development, plasticity, and disease. *Annu. Rev. Cell Dev. Biol.* **24,** 183–209 (2008).
16. Missler, M., Sudhof, T. C. & Biederer, T. Synaptic cell adhesion. *Cold Spring Harb. Perspect. Biol.* **4,** a005694 (2012).
17. Naisbitt, S. *et al.* Shank, a novel family of postsynaptic density proteins that binds to the NMDA receptor/PSD-95/GKAP complex and cortactin. *Neuron* **23,** 569–582 (1999).
18. Sala, C. *et al.* Regulation of dendritic spine morphology and synaptic function by Shank and Homer. *Neuron* **31,** 115–130 (2001).
19. Sheng, M. & Kim, E. The postsynaptic organization of synapses. *Cold Spring Harb. Perspect. Biol.* **3,** a005678 (2011).
20. Luscher, C. & Malenka, R. C. NMDA receptor-dependent long-term potentiation and long-term depression (LTP/LTD). *Cold Spring Harb. Perspect. Biol.* **4,** a005710 (2012).
21. Hardingham, G. E., Arnold, F. J. & Bading, H. A calcium microdomain near NMDA receptors: on switch for ERK-dependent synapse-to-nucleus communication. *Nature Neurosci.* **4,** 565–566 (2001).
22. Malenka, R. C. & Bear, M. F. LTP and LTD: an embarrassment of riches. *Neuron* **44,** 5–21 (2004).
23. Bhakar, A. L., Dolen, G. & Bear, M. F. The pathophysiology of fragile X (and what it teaches us about synapses). *Annu. Rev. Neurosci.* **35,** 417–443 (2012).
24. Bading, H., Ginty, D. D. & Greenberg, M. E. Regulation of gene expression in hippocampal neurons by distinct calcium signaling pathways. *Science* **260,** 181–186 (1993).
25. Dolmetsch, R. E., Pajvani, U., Fife, K., Spotts, J. M. & Greenberg, M. E. Signaling to the nucleus by an L-type calcium channel–calmodulin complex through the MAP kinase pathway. *Science* **294,** 333–339 (2001).
26. Flavell, S. W. *et al.* Activity-dependent regulation of MEF2 transcription factors suppresses excitatory synapse number. *Science* **311,** 1008–1012 (2006).
27. Greenberg, M. E., Ziff, E. B. & Greene, L. A. Stimulation of neuronal acetylcholine receptors induces rapid gene transcription. *Science* **234,** 80–83 (1986).
28. Lin, Y. *et al.* Activity-dependent regulation of inhibitory synapse development by Npas4. *Nature* **455,** 1198–1204 (2008).
29. Flavell, S. W. *et al.* Genome-wide analysis of MEF2 transcriptional program reveals synaptic target genes and neuronal activity-dependent polyadenylation site selection. *Neuron* **60,** 1022–1038 (2008).
30. Hong, E. J., McCord, A. E. & Greenberg, M. E. A biological function for the neuronal activity-dependent component of *Bdnf* transcription in the development of cortical inhibition. *Neuron* **60,** 610–624 (2008).
31. Splawski, I. *et al.* Ca$_V$1.2 calcium channel dysfunction causes a multisystem disorder including arrhythmia and autism. *Cell* **119,** 19–31 (2004).
32. Sanders, S. J. *et al.* Multiple recurrent *de novo* CNVs, including duplications of the 7q11.23 Williams syndrome region, are strongly associated with autism. *Neuron* **70,** 863–885 (2011).
33. Darnell, J. C. *et al.* FMRP stalls ribosomal translocation on mRNAs linked to synaptic function and autism. *Cell* **146,** 247–261 (2011).
34. Morrow, E. M. *et al.* Identifying autism loci and genes by tracing recent shared ancestry. *Science* **321,** 218–223 (2008).
    This paper provides evidence for the hypothesis that genetic mutations associated with ASD may disrupt activity-dependent gene expression programs.
35. Chahrour, M. H. *et al.* Whole-exome sequencing and homozygosity analysis implicate depolarization-regulated neuronal genes in autism. *PLoS Genet.* **8,** e1002635 (2012).
36. Kim, T. K. *et al.* Widespread transcription at neuronal activity-regulated enhancers. *Nature* **465,** 182–187 (2010).
37. Sudhof, T. C. Neuroligins and neurexins link synaptic function to cognitive disease. *Nature* **455,** 903–911 (2008).
38. Tabuchi, K. *et al.* A neuroligin-3 mutation implicated in autism increases inhibitory synaptic transmission in mice. *Science* **318,** 71–76 (2007).
39. Etherton, M. R., Tabuchi, K., Sharma, M., Ko, J. & Sudhof, T. C. An autism-associated point mutation in the neuroligin cytoplasmic tail selectively impairs AMPA receptor-mediated synaptic transmission in hippocampus. *EMBO J.* **30,** 2908–2919 (2011).
40. Etherton, M. R., Blaiss, C. A., Powell, C. M. & Sudhof, T. C. Mouse neurexin-1α deletion causes correlated electrophysiological and behavioral changes consistent with cognitive impairments. *Proc. Natl Acad. Sci. USA* **106,** 17998–18003 (2009).
41. Chubykin, A. A. *et al.* Activity-dependent validation of excitatory versus inhibitory synapses by neuroligin-1 versus neuroligin-2. *Neuron* **54,** 919–931 (2007).
    This study reports the crucial role for neuroligins in activity-dependent synapse maturation and how this pathway is dysregulated by ASD-associated mutations in neuroligin.
42. Kim, J. *et al.* Neuroligin-1 is required for normal expression of LTP and associative fear memory in the amygdala of adult animals. *Proc. Natl Acad. Sci. USA* **105,** 9087–9092 (2008).
43. Choi, Y. B. *et al.* Neurexin-neuroligin transsynaptic interaction mediates learning-related synaptic remodeling and long-term facilitation in *Aplysia*. *Neuron* **70,** 468–481 (2011).
    Using the *Aplysia* model, this study demonstrates that the neurexin–neuroligin interaction regulates activity-dependent synaptic development and plasticity, a process that is dysregulated by an ASD-associated mutation in neuroligin.
44. Thyagarajan, A. & Ting, A. Y. Imaging activity-dependent regulation of neurexin-neuroligin interactions using *trans*-synaptic enzymatic biotinylation. *Cell* **143,** 456–469 (2010).
45. Iijima, T. *et al.* SAM68 regulates neuronal activity-dependent alternative splicing of neurexin-1. *Cell* **147,** 1601–1614 (2011).
46. Sato, D. *et al.* *SHANK1* deletions in males with autism spectrum disorder. *Am. J. Hum. Genet.* **90,** 879–887 (2012).
47. Peca, J. *et al.* *Shank3* mutant mice display autistic-like behaviours and striatal dysfunction. *Nature* **472,** 437–442 (2011).
48. Schmeisser, M. J. *et al.* Autistic-like behaviours and hyperactivity in mice lacking ProSAP1/Shank2. *Nature* **486,** 256–260 (2012).
49. Won, H. *et al.* Autistic-like social behaviour in *Shank2*-mutant mice improved by restoring NMDA receptor function. *Nature* **486,** 261–265 (2012).
    References 47–49 demonstrate the role of Shanks in regulating excitatory neurotransmission.
50. Ehlers, M. D. Activity level controls postsynaptic composition and signaling via the ubiquitin-proteasome system. *Nature Neurosci.* **6,** 231–242 (2003).
51. Bingol, B. & Sheng, M. Deconstruction for reconstruction: the role of proteolysis in neural plasticity and disease. *Neuron* **69,** 22–32 (2011).
52. Hung, A. Y., Sung, C. C., Brito, I. L. & Sheng, M. Degradation of postsynaptic scaffold GKAP and regulation of dendritic spine morphology by the TRIM3 ubiquitin ligase in rat hippocampal neurons. *PLoS ONE* **5,** e9842 (2010).
53. Lee, S. H. *et al.* Synaptic protein degradation underlies destabilization of retrieved fear memory. *Science* **319,** 1253–1256 (2008).
54. Zalfa, F. *et al.* The fragile X syndrome protein FMRP associates with *BC1* RNA and regulates the translation of specific mRNAs at synapses. *Cell* **112,** 317–327 (2003).
55. Huber, K. M., Gallagher, S. M., Warren, S. T. & Bear, M. F. Altered synaptic plasticity in a mouse model of fragile X mental retardation. *Proc. Natl Acad. Sci. USA* **99,** 7746–7750 (2002).
56. Bear, M. F., Huber, K. M. & Warren, S. T. The mGluR theory of fragile X mental retardation. *Trends Neurosci.* **27,** 370–377 (2004).
57. Osterweil, E. K., Krueger, D. D., Reinhold, K. & Bear, M. F. Hypersensitivity to mGluR5 and ERK1/2 leads to excessive protein synthesis in the hippocampus of a mouse model of fragile X syndrome. *J. Neurosci.* **30,** 15616–15627 (2010).
58. Dolen, G. *et al.* Correction of fragile X syndrome in mice. *Neuron* **56,** 955–962 (2007).
59. Park, S. *et al.* Elongation factor 2 and fragile X mental retardation protein control the dynamic translation of Arc/Arg3.1 essential for mGluR-LTD. *Neuron* **59,** 70–83 (2008).
60. Waung, M. W., Pfeiffer, B. E., Nosyreva, E. D., Ronesi, J. A. & Huber, K. M. Rapid translation of Arc/Arg3.1 selectively mediates mGluR-dependent LTD through persistent increases in AMPAR endocytosis rate. *Neuron* **59,** 84–97 (2008).
61. Niere, F., Wilkerson, J. R. & Huber, K. M. Evidence for a fragile X mental retardation protein-mediated translational switch in metabotropic glutamate receptor-triggered Arc translation and long-term depression. *J. Neurosci.* **32,** 5924–5936 (2012).
62. Hou, L. *et al.* Dynamic translational and proteasomal regulation of fragile X mental retardation protein controls mGluR-dependent long-term depression. *Neuron* **51,** 441–454 (2006).
63. Harlow, E. G. *et al.* Critical period plasticity is disrupted in the barrel cortex of *Fmr1* knockout mice. *Neuron* **65,** 385–398 (2010).
64. Pfeiffer, B. E. *et al.* Fragile X mental retardation protein is required for synapse elimination by the activity-dependent transcription factor MEF2. *Neuron* **66,** 191–197 (2010).
65. Barbosa, A. C. *et al.* MEF2C, a transcription factor that facilitates learning and memory by negative regulation of synapse numbers and function. *Proc. Natl Acad. Sci. USA* **105,** 9391–9396 (2008).
66. Pulipparacharuvil, S. *et al.* Cocaine regulates MEF2 to control synaptic and behavioral plasticity. *Neuron* **59,** 621–633 (2008).
67. Han, J. M. & Sahin, M. TSC1/TSC2 signaling in the CNS. *FEBS Lett.* **585,** 973–980 (2011).
68. Kang, H. & Schuman, E. M. A requirement for local protein synthesis in neurotrophin-induced hippocampal synaptic plasticity. *Science* **273,** 1402–1406 (1996).
69. Ehninger, D. *et al.* Reversal of learning deficits in a $Tsc2^{+/-}$ mouse model of tuberous sclerosis. *Nature Med.* **14,** 843–848 (2008).
70. Zhou, J. & Parada, L. F. PTEN signaling in autism spectrum disorders. *Curr. Opin. Neurobiol.* **22,** 873–879 (2012).
71. Auerbach, B. D., Osterweil, E. K. & Bear, M. F. Mutations causing syndromic autism define an axis of synaptic pathophysiology. *Nature* **480,** 63–68 (2011).
    This paper shows that loss of function of either TSC1–TSC2 or FMRP disrupts activity-regulated mRNA translation and synaptic plasticity in opposite directions.
72. Bateup, H. S., Takasaki, K. T., Saulnier, J. L., Denefrio, C. L. & Sabatini, B. L. Loss of Tsc1 *in vivo* impairs hippocampal mGluR-LTD and increases excitatory synaptic function. *J. Neurosci.* **31,** 8862–8869 (2011).
73. Ceman, S. *et al.* Phosphorylation influences the translation state of FMRP-associated polyribosomes. *Hum. Mol. Genet.* **12,** 3295–3305 (2003).
74. Narayanan, U. *et al.* FMRP phosphorylation reveals an immediate-early signaling pathway triggered by group I mGluR and mediated by PP2A. *J. Neurosci.* **27,** 14349–14357 (2007).
75. Muddashetty, R. S. *et al.* Reversible inhibition of PSD-95 mRNA translation by miR-125a, FMRP phosphorylation, and mGluR signaling. *Mol. Cell* **42,** 673–688 (2011).
76. Bhat, S. *et al.* *CACNA1C* (Ca$_v$1.2) in the pathophysiology of psychiatric disease. *Prog. Neurobiol.* **99,** 1–14 (2012).

77. Barrett, C. F. & Tsien, R. W. The Timothy syndrome mutation differentially affects voltage- and calcium-dependent inactivation of $\mathrm{Ca_V}1.2$ L-type calcium channels. *Proc. Natl Acad. Sci. USA* **105,** 2157–2162 (2008).
78. Pasca, S. P. *et al.* Using iPSC-derived neurons to uncover cellular phenotypes associated with Timothy syndrome. *Nature Med.* **17,** 1657–1662 (2011).
79. Skene, P. J. *et al.* Neuronal MeCP2 is expressed at near histone-octamer levels and globally alters the chromatin state. *Mol. Cell* **37,** 457–468 (2010).
80. Cohen, S. *et al.* Genome-wide activity-dependent MeCP2 phosphorylation regulates nervous system development and function. *Neuron* **72,** 72–85 (2011).
This paper demonstrates the importance of activity-dependent regulation of MECP2 in the control of synapse development and behaviour.
81. Chen, W. G. *et al.* Derepression of BDNF transcription involves calcium-dependent phosphorylation of MeCP2. *Science* **302,** 885–889 (2003).
82. Zhou, Z. *et al.* Brain-specific phosphorylation of MeCP2 regulates activity-dependent *Bdnf* transcription, dendritic growth, and spine maturation. *Neuron* **52,** 255–269 (2006).
83. Deng, J. V. *et al.* MeCP2 in the nucleus accumbens contributes to neural and behavioral responses to psychostimulants. *Nature Neurosci.* **13,** 1128–1136 (2010).
84. Li, H., Zhong, X., Chau, K. F., Williams, E. C. & Chang, Q. Loss of activity-induced phosphorylation of MeCP2 enhances synaptogenesis, LTP and spatial memory. *Nature Neurosci.* **14,** 1001–1008 (2011).
85. Noutel, J., Hong, Y. K., Leu, B., Kang, E. & Chen, C. Experience-dependent retinogeniculate synapse remodeling is abnormal in MeCP2-deficient mice. *Neuron* **70,** 35–42 (2011).
86. Dani, V. S. *et al.* Reduced cortical activity due to a shift in the balance between excitation and inhibition in a mouse model of Rett syndrome. *Proc. Natl Acad. Sci. USA* **102,** 12560–12565 (2005).
87. Chao, H. T., Zoghbi, H. Y. & Rosenmund, C. MeCP2 controls excitatory synaptic strength by regulating glutamatergic synapse number. *Neuron* **56,** 58–65 (2007).
88. Moretti, P. *et al.* Learning and memory and synaptic plasticity are impaired in a mouse model of Rett syndrome. *J. Neurosci.* **26,** 319–327 (2006).
89. Greer, P. L. *et al.* The Angelman syndrome protein Ube3A regulates synapse development by ubiquitinating Arc. *Cell* **140,** 704–716 (2010).
This study provides insight into the molecular mechanisms underlying Angelman syndrome and the role of neuronal activity in regulating UBE3A and its targets.
90. Jiang, Y. H. *et al.* Mutation of the Angelman ubiquitin ligase in mice causes increased cytoplasmic p53 and deficits of contextual learning and long-term potentiation. *Neuron* **21,** 799–811 (1998).
91. Yashiro, K. *et al.* Ube3a is required for experience-dependent maturation of the neocortex. *Nature Neurosci.* **12,** 777–783 (2009).
92. Wallace, M. L., Burette, A. C., Weinberg, R. J. & Philpot, B. D. Maternal loss of Ube3a produces an excitatory/inhibitory imbalance through neuron type-specific synaptic defects. *Neuron* **74,** 793–800 (2012).
93. Kelleher, R. J. III *et al.* High-throughput sequencing of mGluR signaling pathway genes reveals enrichment of rare variants in autism. *PLoS ONE* **7,** e35003 (2012).
94. Pinto, D. *et al.* Functional impact of global rare copy number variation in autism spectrum disorders. *Nature* **466,** 368–372 (2010).
95. Kelleher, R. J. III, Govindarajan, A., Jung, H. Y., Kang, H. & Tonegawa, S. Translational control by MAPK signaling in long-term synaptic plasticity and memory. *Cell* **116,** 467–479 (2004).
96. Guy, J., Gan, J., Selfridge, J., Cobb, S. & Bird, A. Reversal of neurological defects in a mouse model of Rett syndrome. *Science* **315,** 1143–1147 (2007).
97. Giacometti, E., Luikenhuis, S., Beard, C. & Jaenisch, R. Partial rescue of MeCP2 deficiency by postnatal activation of MeCP2. *Proc. Natl Acad. Sci. USA* **104,** 1931–1936 (2007).
98. van Woerden, G. M. *et al.* Rescue of neurological deficits in a mouse model for Angelman syndrome by reduction of αCaMKII inhibitory phosphorylation. *Nature Neurosci.* **10,** 280–282 (2007).

Acknowledgements M.E.G. is supported by NIH grant RO1NS048276 and the Rett Syndrome Research Trust. D.H.E. is supported by NIH grant K08MH90306.

Author Information Reprints and permissions information is available at www.nature.com/reprints. The authors declare no competing financial interests. Readers are welcome to comment on the online version of this article at go.nature.com/trarbs. Correspondence should be addressed to M.E.G. (michael_greenberg@hms.harvard.edu).

doi:10.1038/nature11861

mTOR is a key modulator of ageing and age-related disease

Simon C. Johnson[1], Peter S. Rabinovitch[1] & Matt Kaeberlein[1,2]

Many experts in the biology of ageing believe that pharmacological interventions to slow ageing are a matter of 'when' rather than 'if'. A leading target for such interventions is the nutrient response pathway defined by the mechanistic target of rapamycin (mTOR). Inhibition of this pathway extends lifespan in model organisms and confers protection against a growing list of age-related pathologies. Characterized inhibitors of this pathway are already clinically approved, and others are under development. Although adverse side effects currently preclude use in otherwise healthy individuals, drugs that target the mTOR pathway could one day become widely used to slow ageing and reduce age-related pathologies in humans.

The mechanistic target of rapamycin (mTOR) story began in the 1970s when new antifungal activity was discovered in soil samples from the Polynesian island of Rapa Nui (Box 1). Thus, the compound, which was isolated from *Streptomyces hygroscopicus*, was named rapamycin. Also known as sirolimus, rapamycin was widely studied as an immunosuppressant before its mechanism of action was well understood, and in 1999 it was approved for use in post-transplantation therapy. Since then, rapamycin and several derivative compounds (including everolimus, temsirolimus, ridaforolimus, umirolimus and zotarolimus, collectively referred to as 'rapamycins' in this Review) have been approved for a variety of uses, including prevention of restenosis following angioplasty and as a treatment for certain forms of cancer.

Studies in the budding yeast *Saccharomyces cerevisiae* first identified the target of rapamycin genes *TOR1* and *TOR2* as genetic mediators of rapamycin's growth inhibitory effects, and soon afterwards the mTOR protein was purified from mammalian cells and demonstrated to be the physical target of rapamycin[1]. mTOR is a serine/threonine protein kinase of the phosphatidylinositol-3-OH kinase (PI(3)K)-related family that functions as a master regulator of cellular growth and metabolism in response to nutrient and hormonal cues[2]. mTOR functions in two distinct complexes: mTOR complex 1 (mTORC1) and mTOR complex 2 (mTORC2) (Fig. 1). Rapamycins inhibit mTORC1 by binding the FK506-binding protein FKBP12, which then interacts physically with the complex and decreases activity. Although mTORC2 is not directly affected by rapamycin, chronic exposure can sequester mTOR from mTORC2, inhibiting mTORC2 assembly. This effect on mTORC2 is thought to contribute to metabolic complications associated with chronic rapamycin treatment, including glucose intolerance and abnormal lipid profiles[3] (described further later).

Much more is known about both the upstream regulation and downstream outputs of mTORC1 compared with mTORC2. mTORC1 is activated by insulin and other growth factors through PI(3)K and AKT kinase signalling[1]. mTORC1 is also activated by environmental nutrients (for example, amino acids) and repressed by AMP-activated protein kinase (AMPK), a key sensor of cellular energy status (discussed further later). In response to these growth signals, mTORC1 is thought to promote messenger RNA translation and protein synthesis through at least two mTORC1 substrates, ribosomal protein S6 kinases (S6Ks) and eukaryotic translation initiation factor 4E-binding protein 1 (4E-BP1). mTORC1 also promotes lipid biosynthesis, represses degradation through the autophagy pathway, and regulates glucose metabolism and mitochondrial function through the hypoxic response transcription factor HIF-1α (Fig. 1) as well as the peroxisome-proliferator-activated receptor coactivator PGC-1α. These multiple inputs and outputs place mTORC1 as a key regulatory nexus, modulating anabolic processes versus catabolic processes in response to nutrients, growth cues and cellular energy status.

mTOR and known longevity factors

The first indication that mTOR regulates ageing came from studies in *S. cerevisiae*, in which it was found that deletion of the gene encoding the yeast orthologue of S6K — *SCH9* — resulted in a doubling of yeast chronological lifespan (defined as the duration of time that cells in stationary phase remain viable)[4]. Soon after, mTORC1 was shown to affect longevity with the finding that mutation or RNA interference (RNAi) knockdown of mTOR (*let-363*) or the mTORC1 component raptor (*daf-15*) could extend lifespan in the nematode *Caenorhabditis elegans*[5,6]. This was followed by studies in the fruitfly *Drosophila melanogaster* and using the yeast replicative ageing model, showing that mutations in mTOR and several other components of the mTORC1 pathway also increase lifespan in these systems[7,8]. In addition, a series of studies showed that rapamycin extended lifespan in yeast[9,10], nematodes[11], fruitflies[12] and mice[13–15], firmly establishing mTORC1 as a central, evolutionarily conserved regulator of longevity (Table 1).

Dietary restriction

Reduction in nutrient intake in the absence of malnutrition, or dietary restriction (also referred to as calorie restriction), extends lifespan in many different species[16]. In fact, other than mTORC1 inhibition, dietary restriction is currently the only intervention known to extend lifespan in yeast ageing models and in worms, flies and mice. mTORC1 is thought to play a part in mediating longevity and health benefits as a result of dietary restriction, which is intuitive given its function in responding to nutrient and growth cues. Dietary restriction reduces mTORC1 activity in invertebrate organisms and in some mammalian tissues, and pharmacological or genetic disruption of mTORC1 is sufficient to extend lifespan in both invertebrates and mice under non-dietary restriction conditions[2].

Genetic evidence that suggests mTORC1 acts downstream of dietary restriction comes mainly from epistasis experiments in which

---

[1]Department of Pathology, University of Washington, Seattle, Washington 98195, USA. [2]Institute of Aging Research, Guangdong Medical College, Dongguan 523808, China.

dietary restriction is combined with mutations in the mTORC1 pathway. In yeast, for example, dietary restriction fails to further extend replicative lifespan when combined with deletion of the genes encoding the mTOR and S6K homologues[8]. A similar relationship is observed when RNAi knockdown of mTOR is combined with dietary restriction in *C. elegans*[17]. In fruitflies, 4E-BP1 is required for maximal lifespan extension from at least one method of dietary restriction[18], and lifespan extension from dominant-negative alleles of mTOR and S6K extend lifespan in fruitflies in a nutrient-dependent manner[7]. However, the interaction between dietary restriction and mTORC1 signalling is complex, as evidenced in reports that RNAi knockdown of S6K and translation initiation factors yield an additive lifespan extension when combined with dietary restriction in *C. elegans*[17]. Interpretation of these types of experiments is complicated by the fact that several different dietary restriction protocols are used, and quantitative assessments of activity are often missing. Despite these challenges, numerous studies have linked mTORC1-regulated processes, including both reduced mRNA translation and induction of autophagy, to lifespan extension from dietary restriction in different organisms (discussed further later). Thus, there is broad consensus that reduced mTORC1 signalling probably contributes to longevity and health benefits as a result of dietary restriction[19].

Insulin/IGF-1-like signalling.

Reduced insulin/IGF-1-like signalling (IIS) increases longevity in nematodes, fruitflies and mice[16,19]. Lifespan extension from reduced IIS is mediated by the FOXO family of transcription factors, and several studies have associated polymorphisms in FOXO3 with longevity in humans[19]. mTOR activity is linked to IIS through multiple connections (Fig. 2). As already mentioned, mTORC1 is activated by IIS through AKT, and mTORC1 can negatively regulate IIS through S6K, which inhibits insulin receptor substrate 1 (IRS-1)[20]. Additional studies have indicated that FOXO3A can transcriptionally regulate tuberous sclerosis protein 1 (TSC1), an upstream inhibitor of mTORC1, in mammalian cells[21] and that the mTORC1 target 4E-BP1 is also a transcriptional target of FOXO in fruitflies[22]. mTOR and IIS also interact through mTORC2, which activates AKT to repress FOXO1 and FOXO3 in mammalian cells[23].

Genetic studies that examine the relationship between mTORC1 and IIS with respect to longevity reflect this complexity. In nematodes, the FOXO orthologue *daf-16* is proposed to repress raptor expression, and lifespan extension from mutation of raptor requires DAF-16 — a defining characteristic of components of the IIS pathway[5]. Paradoxically, mutation of mTOR itself[6] or treatment with rapamycin[11] does not require DAF-16 for lifespan extension. Similarly, lifespan extension that results from mutations or RNAi knockdown of mTOR-pathway downstream targets such as S6K and the translation initiation factors eIF4E and eIF4G does not require DAF-16 (refs 17, 24). Thus, current evidence supports the idea that mTORC1 modulates ageing by mechanisms that overlap but

BOX 1

Key events for rapamycins in ageing

- **1970** Soil samples containing rapamycin-producing *Streptomyces hygroscopicus* taken from the Polynesian island of Rapa Nui
- **1975** Purification of rapamycin and identification of fungicidal activity
- **1977** Immunosuppressive activity discovered
- **1984** Antitumour activity discovered
- **1991** *TOR1* and *TOR2* genes identified in yeast
- **1994** mTOR gene *FRAP1*, now *MTOR*, identified in mammals
- **1995** Mechanism of action of rapamycin discovered
- **1999** Approved by US Food and Drug Administration (FDA) for use in preventing host-rejection in patients undergoing kidney transplantation
- **2003** Approved by FDA for use in drug-eluting stents
- **2006** Shown to extend lifespan in budding yeast
- **2007** Approved by FDA for treatment of renal-cell carcinoma
- **2008** Approved by FDA for treatment of mantle cell lymphoma
- **2009** Shown to extend lifespan in mice
- **2010** Shown to extend lifespan in fruitflies
- **2010** Approved by FDA for treatment of tuberous sclerosis
- **2011** Shown to improve outcome in mouse models of Alzheimer's disease
- **2011** Approved by FDA for treatment of pancreatic cancer
- **2012** Shown to extend lifespan in nematodes
- **2012** More than 1,300 clinical trials under way or completed

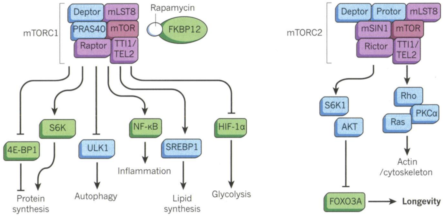


Figure 1 | The two mTOR complexes have distinct constituent proteins and regulate different downstream processes. Here (figure represents data from studies in mice) mTORC1 comprises deptor, PRAS40, raptor, mLST8, mTOR and TTI1–TEL2. mTORC2 is comprised of deptor, mLST8, protor, rictor, mSIN1, mTOR and TTI1–TEL2. Rapamycin binds to FKBP12 and inhibits mTORC1 by disrupting the interaction between mTOR and raptor. Regulation of lipid synthesis by mTORC1 is thought to occur mainly through sterol-regulatory-element-binding protein transcription factors (shown here as SREBP1) by a mechanism that is not completely understood. mTORC1 negatively regulates autophagy through multiple inputs, including inhibitory phosphorylation of ULK1, preventing formation of the ULK1–ATG13–FIP200 complex (which is required for initiation of autophagy). mTORC1 promotes protein synthesis through activation of the translation initiation promoter S6K and through inhibition of the inhibitory mRNA cap binding 4E-BP1, and regulates glycolysis through HIF-1α. mTORC2 inhibits FOXO3a through S6K1 and AKT, which can lead to increased longevity. The complex also regulates actin cytoskeleton assembly through protein kinase C α (PKCα), Rho GTPases and Ras proteins.

Table 1 | Comparison of species in which genetic or pharmacological inhibition of an mTORC1-pathway component extends lifespan

	*Saccharomyces cerevisiae**	*Caenorhabditis elegans*	*Drosophila melanogaster*	*Mus musculus*
Rapamycin	Yes[9,10]	Yes[11]	Yes[12]	Yes[13–15]
mTOR gene mutation and knockdown	Yes[8,9]	Yes[6]	Yes[7]	Yes[3]
Raptor gene mutation and knockdown	Not reported	Yes[5]	Not reported	Not reported
Tsc1 and *Tsc2* activation	Not applicable†	Not applicable	Yes[7]	Not reported
S6K gene mutation and knockdown	Yes[4,8]	Yes[24]	Yes[7]	Yes[37]
4E-BP activation	Not applicable	Not applicable	Yes[18]	Not reported
Translation initiation factor mutation and knockdown	Yes[44]	Yes[17,24,74,75]	Not reported	Not reported
Ribosomal protein mutation and knockdown	Yes[44,47,76]	Yes[17,75]	Not reported	Not reported

*Includes both replicative and chronological lifespan; †Not applicable is used in cases for which homologues have not yet been identified.

are distinct from IIS. This is consistent with the model that mTORC1 acts mainly downstream of dietary restriction, which, based on similar epistasis experiments, also seems to increase lifespan by mechanisms that are overlapping but distinct from IIS.

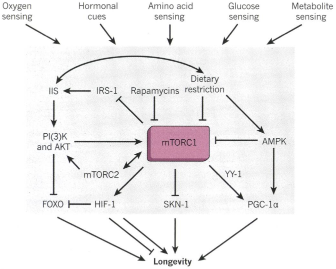


Figure 2 | Interactions between mTOR and other longevity pathways. mTORC1 responds to a variety of environmental cues and communicates with several known longevity factors in a complex network of interactions. mTORC1 activity limits longevity, and several pathways related to mTORC1 can result in the extension of lifespan. Rapamycins inhibit mTORC1 and decrease its activity. mTORC1 can be activated by environmental nutrients such as amino acids. Hormonal cues can stimulate IIS, which can increase longevity through PI(3)K and AKT mediation of the FOXO family of transcription factors. mTORC1 can also negatively regulate IIS through inhibition of IRS-1. Sensing of low oxygen levels stimulates mTORC1 to activate the hypoxic response by enhancing translation of HIF-1, which inhibits FOXO family members and increases longevity. HIF-1 is also thought to extend longevity at high and low temperatures, or inhibit it at low temperatures. Inhibition of mTORC1 signalling has been linked to stress resistance, and inhibition of the stress response transcription factor SKN-1 (an orthologue of Nrf2) by mTORC1 has been implicated in lifespan extension. Glucose and metabolite sensing as a result of dietary restriction can lower mTORC1 signalling partly through activation of AMPK, which can inhibit mTORC1 and promote longevity. The complex actively promotes mitochondrial biogenesis and metabolism through PGC-1α and YY-1. Inhibition of this results in greater mitochondrial respiration, leading to increased longevity. Dietary restriction and IIS can also influence each other and their related pathways. The relationship between mTORC2 and mTORC1 is still unclear. The interactions shown have been described in studies of yeast, nematodes, fruitflies and mice.

AMP-activated protein kinase

AMPK is a conserved sensor of energy status that is activated in response to low ATP levels and negatively regulates mTORC1 (Fig. 2). Overexpression of AMPK is sufficient to extend lifespan in *C. elegans*, and the AMPK activator metformin (a widely used antidiabetic drug) is reported to extend lifespan in *C. elegans* and in short-lived, cancer-prone strains of mice[25–27]. AMPK inhibits mTORC1 through at least two distinct mechanisms: AMPK phosphorylates TSC2 on conserved serine sites, resulting in activation of TSC2 and downregulation of mTORC1 activity[28], and AMPK directly phosphorylates raptor to impair mTORC1 signalling[29]. These data suggest that dietary restriction could reduce mTORC1 signalling partly through activation of AMPK. The actual situation is likely to be more complex, however, as both AMPK and mTOR interact with multiple additional factors, and it has been reported that activation of AMPK extends the lifespan of *C. elegans* through the CREB-regulated transcriptional coactivator 1 (CRTC-1) rather than through mTORC1 (ref. 30). Importantly, so far, there is no direct evidence for regulation of mTORC1 by AMPK in *C. elegans*, which does not have a TSC2 orthologue.

The hypoxic response

HIF-1α is a highly conserved transcription factor that is activated in response to low oxygen availability (hypoxia) and regulates the expression of a suite of genes involved in maintaining cellular homeostasis under these conditions[31]. Under normoxic conditions, HIF-1α is targeted for proteasomal degradation, and ectopic stabilization of HIF-1α is a hallmark of many cancers. A major feature of the hypoxic response is to drive ATP production under oxygen-limiting conditions by promoting a metabolic shift from oxidative phosphorylation towards glycolysis and lactic acid fermentation. mTORC1 activates the early hypoxic response in mammals by enhancing translation and stabilization of HIF-1 itself, as well as by enhancing translation of mRNAs encoding a subset of HIF-1 target genes that include vascular endothelial growth factor (VEGF)[32]. Prolonged hypoxia, however, results in downregulation of mTORC1.

The hypoxic response has been implicated in ageing by studies in *C. elegans* in which HIF-1 acts as both a positive and negative regulator of longevity[31]. Stabilization of HIF-1 results in lifespan extension by a mechanism that is distinct from both IIS and dietary restriction[33,34]. Interestingly, deletion of *hif-1* also extends lifespan of animals grown at a high temperature (25 °C), but not at a low temperature (15 °C)[35]. In this context, HIF-1 has been proposed to act downstream of dietary restriction and mTORC1 (ref. 36) (Fig. 2). However, direct evidence for regulation of HIF-1 by mTORC1 in *C. elegans* has yet to be obtained.

mTOR in mammalian ageing and healthspan

The first indication that mTOR modulates mammalian ageing came from longevity studies as part of the National Institute on Aging's Interventions Testing Program (ITP). Rapamycin was found to significantly extend lifespan in a genetically heterogeneous strain background at three independent test locations[13]. A particularly interesting aspect of this study was that the animals did not begin receiving the drug until they had reached 600 days of age, roughly equivalent to 60 years of age in a person. Follow-up studies that began rapamycin treatment at an earlier age replicated the initial result but failed to show substantially larger effects on longevity[14]. In mice treated with rapamycin from 6 months of age, mean lifespan extension was about 18% in females, but only 10% in males[14]. Both higher and lower doses of rapamycin are currently being tested, and it will be of particular interest to see whether a greater longevity benefit can be achieved. Soon after the ITP rapamycin study was published, knockout of the mouse gene encoding S6K1 (*Rps6kb1*) was reported to increase lifespan in females[37]. In this case, the lifespan extension was sex-specific, with males receiving no longevity benefit.

A key prediction of the hypothesis that ageing is caused by specific molecular changes is that it should be possible to slow the ageing process and, thereby, delay the onset and progression of multiple age-related diseases (Fig. 3). Thus, slowing ageing should increase both lifespan and healthspan — the period of life spent in relatively good health, free from chronic disease or disability. This seems to be the case for dietary restriction, which not only extends lifespan but also delays the incidence of age-related decline and disease in rodents, including cancer, cognitive decline and neurodegeneration[16]. Similarly, dietary restriction reduces the incidence of cancer, cardiovascular disease, brain atrophy and diabetes in rhesus monkeys[38]. Emerging evidence suggests that, like dietary restriction, mTORC1 inhibition may have similar positive effects on multiple age-related pathologies in rodents and, in some cases, humans (Box 2).

From a public-health perspective, an intervention that results in compression of morbidity, in which most of a lifetime's illness is compressed into a shorter period of time near the end of life[39], is particularly desirable. Although it is too early to know whether rapamycins could represent such an intervention, end-of-life analysis of the long-lived rapamycin-treated mice in the ITP studies indicates that the spectrum of causes of death was not significantly altered[14]. However, age-related alterations in heart, liver, adrenal glands, endometrium and tendon, as well as the decline in spontaneous activity, all occur more slowly in rapamycin-treated mice[40]. Another study found that, in addition to lifespan extension, spontaneous tumorigenesis and deaths owing to cancer in the inbred C57BL/6 mouse strain were significantly reduced by rapamycin[15]. These data suggest that rapamycin slows the ageing process in mice such that many normal causes of morbidity and death are delayed. The evidence that rapamycin can protect against multiple age-related diseases (Box 2) — including reduced cognitive function[41,42] — is consistent with a compression of morbidity, although further studies will be needed to definitively test this. As previously alluded to, there has not yet been a careful analysis of different rapamycin regimens, and the observed effects on longevity and healthspan may be quite different at higher or lower doses of the drug.

Mechanisms of longevity regulation by mTOR

In recent years, much effort has been applied to defining the mechanisms by which inhibition of mTORC1 enhances longevity. Not surprisingly, the picture that has emerged is complex. Multiple mTORC1-regulated processes seem to contribute to the pro-longevity effects of mTORC1 inhibition in a coordinated and overlapping manner.

mRNA translation

Among the most crucial of mTORC1 functions is the promotion of mRNA translation and protein synthesis under conditions that are favourable for growth, and there is evidence that regulation of mRNA translation can modulate longevity in yeast, nematodes, fruitflies and mice[43]. For example, mutations in S6K have been found to extend lifespan in all four of these model organisms (Table 1). Likewise, mutation or knockdown of multiple translation initiation factors and ribosomal proteins has also been shown to extend lifespan in yeast and nematodes[44], and overexpression of an activated allele of 4E-BP extends lifespan in flies[18].

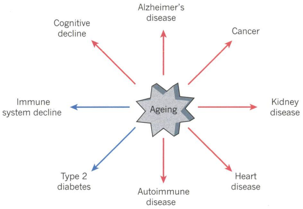


Figure 3 | The impact of mTORC1 on diseases of ageing. Ageing drives the onset and progression of multiple disorders that are modulated by mTORC1 signalling. Data from animal and human studies indicate that some disorders (red arrows) are improved by rapamycins. However, for others (blue arrows), although influenced by rapamycins, evidence suggests there are both beneficial and detrimental consequences of mTORC1 inhibition.

It has been proposed that a global reduction in mRNA translation could be beneficial during ageing by allowing endogenous protein repair and degradation machinery to better maintain protein homeostasis in the face of protein aggregation and oxidative damage[45]. Alternatively, a general reduction in mRNA translation could attenuate age-associated pathologies that result from 'hyperfunctional' overactive biosynthetic and proliferative processes that are important during development but detrimental during adulthood[46]. Instead of a global reduction in mRNA translation, however, molecular evidence suggests that differential translation of specific mRNAs is more important for effects on lifespan. The evidence for this came from yeast, in which lifespan extension in ribosomal protein mutants is associated with increased translation of the transcriptional activator Gcn4, which regulates the expression of many target genes, including several that are involved in enhanced stress resistance and response to amino acid starvation[47]. This regulation of Gcn4 is mediated by structural features of the 5′ region of the *GCN4* mRNA that cause it to be preferentially translated when mTORC1 is inhibited. Mechanistically similar observations have been made in both nematodes and fruitflies, in which a subset of mRNAs are translated more efficiently when global mRNA translation is reduced through inhibition of mTORC1-pathway components. In worms, these differentially translated mRNAs are enriched for stress response genes, and in flies they are enriched for mRNAs encoding components of the mitochondrial electron transport chain[18,48]. Although the specific mRNAs seem to differ across species, a common theme of enhanced translation efficiency of metabolic and stress genes in response to mTORC1 inhibition has emerged.

The mechanism by which mTORC1 regulates mRNA translation in mammals continues to be an active area of research. Although *S6k1* knockout mice are long-lived and have a small body size[37], there is no direct evidence that these animals have reduced rates of global mRNA translation. In both yeast and worms, a deficiency of S6K causes a substantial reduction in global mRNA translation and protein synthesis[24,47], so it would be somewhat surprising if mice that lack S6K1 do not. However, one recent study has questioned whether S6K1 is a major component of mTORC1 regulation of mRNA translation in mammalian cells[49]. In experiments using the mTOR catalytic inhibitor Torin 1, acute control of translation by mTORC1 in $p53^{-/-}$ mouse embryonic fibroblasts was found to be mediated mainly by 4E-BPs, and was largely limited to mRNAs with 5′ terminal oligopyrimidine (TOP) motifs[49]. This study also found no evidence that complexity of the 5′ untranslated region is a determinant of mTORC1-dependent differential mRNA translation[49],

BOX 2

Effects of mTORC1 inhibition on age-related disease

Inhibition of mTORC1 has positive, and sometimes negative, effects on the age-related pathologies in model organisms and, in some cases, humans.

- **Neurodegenerative diseases** Protection against dopaminergic neuron death in fruitfly and mouse models of Parkinson's disease[77,78], delayed disease progression in mouse models of Alzheimer's disease[79,80] and frontotemporal lobar dementia[81], and improved spatial learning and memory in old-aged mice[41,42].
- **Cancer** Reduced growth of solid-tumour cell lines, but disappointing efficacy in clinical trials, with the exception of renal-cell carcinoma, breast cancer and several rare forms of cancer[82].
- **Heart disease** Widespread use to prevent restenosis following angioplasty. Reduced hypertrophy[83] and regression of failure[84] in mouse aortic-constriction model. Reduced cardiomyopathy in models of the zebrafish *Danio rerio*[85] and in mice[86].
- **Metabolic disease and obesity** Resistance to diet-induced[60,87] and age-associated[88] obesity in mice. However, inhibition is also associated with glucose intolerance and insulin resistance in mice[3], and dyslipidaemia in humans.
- **Immune function** Rapamycins are used in combined immunosuppression therapies in humans. Rapamycin pre-treatment enhances immune function in tuberculosis[89], influenza[68] and antitumour vaccine responses in mice[90] and in vaccinia vaccine responses in non-human primates[91].
- **Kidney disease** Decreased rejection and nephrotoxicity in allograft and renal-cancer therapies in patients[92]. Improved outcome in animal models of polycystic kidney disease and diabetic nephropathy.
- **Other health effects** Beneficial effects in animal models of age-associated retinopathy[93], and in a variety of disorders that are not obviously age-related, including cerebral ischaemia and reperfusion injury[94], depression[41], autism spectrum disorder[95], muscular dystrophy[86], epilepsy[96], food allergies[97] and autoimmune disorders[98]. Decreased need for anti-VEGF intravitreal injections in humans with age-related macular degeneration[99]. Reduced toxicity and enhanced clearance of progerin in cells from patients with Hutchinson–Gilford progeria syndrome[100]. Negative effects include testicular degeneration and cataracts in mice[58], oral mucositis and stomatitis in humans, and impaired wound healing in humans and animal models.

which is in contrast with the yeast and invertebrate models already described. It may be that the precise mechanisms by which mTORC1 regulates mRNA translation, and the impact of this regulation on longevity, is different in mammals and invertebrate species. Alternatively, these differences could arise from effects of the catalytic inhibition of mTOR, as opposed to genetic inhibition or inhibition with rapamycin, or from differences in mTOR function in cells grown between culture and *in vivo*.

Autophagy

Activation of autophagy is another key mTORC1-regulated process that probably has a central role in promoting longevity. Autophagy is a major degradation pathway in eukaryotic cells that is essential for removing damaged organelles and macromolecules from the cytoplasm and recycling amino acids during periods of starvation[50]. Evidence suggests that autophagic degradation declines with age, and it has been proposed that this leads to an accumulation of damage, such as protein aggregates and degenerate mitochondria, that contribute to age-related cellular dysfunction[51]. Activation of autophagy by inhibition of mTORC1 presumably maintains cellular function during ageing by allowing enhanced degradation of aged cellular components. However, this hypothesis has been difficult to test directly, partly because the tools for quantifying autophagic flux are inadequate and rely mainly on secondary assays — such as the abundance of proteins involved in various steps of autophagy. Evidence from studies in yeast and invertebrates supports the model that mTORC1-mediated induction of autophagy is required for lifespan extension from dietary restriction or from rapamycin[12,52,53], although it remains to be determined whether induction of autophagy is sufficient to promote longevity in the absence of mTORC1 inhibition. In addition to longevity, aberrant regulation of autophagy has been linked to several diseases of ageing, including cancer, diabetes, cardiovascular disease and neurodegenerative diseases[50], and it seems likely that enhanced autophagy underlies many of the beneficial effects of mTORC1 inhibition in these disease models.

Stress resistance and xenobiotic metabolism

Inhibition of mTORC1 signalling has been linked most clearly to enhanced stress resistance in yeast and nematodes. In yeast, for example, reduced activity of Tor1 and Sch9 activates the protein kinase Rim15, which in turn activates the stress response transcription factors Msn2, Msn4 and Gis1 (ref. 54). Induction of these stress response pathways is required for yeast chronological lifespan extension from dietary restriction, or in response to reduced mTORC1 signalling through deletion of *SCH9* (ref. 54). In nematodes, treatment with rapamycin has been shown to activate the stress response transcription factor SKN-1, which is encoded by the orthologue of mammalian *NRF2* (ref. 11). SKN-1 is required for lifespan extension from rapamycin[11], and has also been implicated in lifespan extension from dietary restriction[55]. Interestingly, NRF2 is activated in mice treated with rapamycin and in those subjected to dietary restriction[56]. In both cases, NRF2 activation results in the induction of several enzymes involved in xenobiotic defence, suggesting that enhanced resistance to such environmental insults may be important for lifespan extension.

Mitochondrial function

Regulation of mitochondrial function by mTORC1 is complex and seems to involve multiple mechanisms. As mentioned previously, mTORC1 promotes activation of HIF-1, which in turn enhances glycolytic flux while simultaneously downregulating mitochondrial oxygen consumption[32]. Another report suggests that mTORC1 actively promotes mitochondrial biogenesis and metabolism through PGC-1α and the transcription factor YY-1 (ref. 57). The best evidence for a direct mitochondrial role in longevity downstream of mTOR signalling comes from yeast, in which inhibition of mTORC1 results in a metabolic shift towards greater mitochondrial respiration, thereby increasing chronological lifespan[58]. This effect has been proposed to involve an adaptive signalling response owing to elevated levels of mitochondrial superoxide[59]. Mice that lack mTORC1 activity in adipose tissue also show enhanced mitochondrial respiration[60], suggesting a similar relationship between mTORC1 and mitochondrial function in mammals, at least in some tissues.

Inflammation

Inflammation is associated with several age-related disorders[61], and a reduction in inflammation is proposed to be a primary mechanism by which dietary restriction promotes longevity and healthspan[62]. Hyperactivation of mTOR is often associated with inflammation, and rapamycin has been shown to have anti-inflammatory effects in multiple settings, including chronic kidney disease[63], vascular inflammation after angioplasty[64], atherosclerotic plaques[65] and lung infection[66]. Thus, a reduction in chronic, age-associated inflammation is another attractive mechanism by which mTORC1 inhibition could slow multiple age-related pathologies in mammals.

Stem cells

In mammals, a decline in stem-cell function is likely to be an important cause of age-related pathology[67]. There is increasing evidence that mTORC1 has a central role in this process, and that inhibition of mTORC1 can preserve, and perhaps even rejuvenate, stem-cell function in a variety of tissues. For example, rapamycin protects old mice from an immune challenge with influenza virus, which has been attributed to a rejuvenation of haematopoietic stem-cell function in the treated animals[68]. Furthermore, treating old mice with rapamycin has been found to enhance intestinal stem-cell function[69]. Interestingly, the mechanism accounting for this improvement is linked to inhibition of mTORC1 in the surrounding Paneth cells, resulting in a more favourable niche, rather than a direct effect of mTORC1 inhibition in the intestinal stem cells. In another study, dietary restriction was shown to enhance the function of muscle stem cells in both young and old animals through both cell-intrinsic and cell-extrinsic mechanisms[70]. Although the authors did not directly examine a role for mTORC1, these effects have been proposed to be partly mediated by reduced mTORC1 activity[71].

Future directions and perspectives

mTOR is a key modulator of ageing in evolutionarily divergent organisms, ranging from yeast to rodents, and it is likely that this function has been conserved to some extent in humans. The complexity of the mTOR network presents a hurdle in defining the mechanistic details of how mTOR influences longevity and healthspan. Several mTOR-regulated processes are likely to contribute to the effects of mTORC1 inhibition on ageing and disease, and it will be challenging to untangle the context-specific and tissue-specific relationships. Nonetheless, impressive progress has been made in this area over the past few years, and we anticipate that this trend will continue.

As the list of beneficial effects of rapamycin in invertebrate and mouse models continues to grow longer (Box 2), it becomes increasingly tempting to speculate on similar benefits in people. As mentioned previously, rapamycins have been approved clinically for a variety of uses but have yet to be tested against a broad spectrum of age-related diseases. At the time of writing, the search term 'rapamycin' generates 1,343 results in the National Institutes of Health clinical trials database (http://clinicaltrials.gov). We therefore anticipate a bounty of additional data on the effects of rapamycin in humans over the next few years.

All of this then raises the question, at what point should we consider giving these drugs to otherwise healthy people? Perhaps the greatest upside of mTORC1 inhibitors is their potential to delay cognitive decline during ageing. The evidence for improved cognitive function in old mice treated with rapamycin is striking[41]. Current estimates suggest cognitive decline can be detected as early as 45 years of age in otherwise healthy people[72]. Loss of cognitive function is a leading concern among geriatricians and their elderly patients and is a significant and growing public-health burden. If mTORC1 inhibition has even a modest positive effect on cognitive function, it could improve the quality of life for millions of middle-aged and older adults. Add to this the likelihood that risks of developing some forms of cancer, cardiovascular disease, and neurodegenerative disease would be reduced, then mTOR inhibitors may offer an attractive opportunity to have a significant impact on preventive health care.

Before this can happen, however, important questions must be answered. Rapamycin is not without side effects, including hyperlipidaemia and hyperglycaemia, anaemia and stomatitis, in patients. A recent study of long-term rapamycin treatment in mice reported increased incidence of cataracts and testicular degeneration[40]. The effects of mTORC1 inhibition on immune function and wound healing are also of particular concern. Although it is unclear whether rapamycin alone has substantial immunosuppressive effects in healthy individuals, it undoubtedly is immunomodulatory. It would be unfortunate to take a drug that slows the rate of ageing, only to succumb to infection from an otherwise innocuous bacterium or virus at an early age. Optimal dosage and duration of treatment are also unknown factors. Almost no information exists on the most effective dose of rapamycin for longevity or healthspan in mice, and the current data suggest that rapamycin therapy that begins late in life is nearly as effective as therapy that begins early in life, at least for longevity. Understanding whether this is also true for a variety of age-related pathologies will be important. Assessing risks and side effects from studies of patients with disorders such as the rare lung disease lymphangioleiomyomatosis, for which rapamycin is indicated for chronic use, may be informative in determining the clinical potential of this drug and other mTOR inhibitors in the setting of age-related disease.

The development of chemical inhibitors of mTOR, as well as drugs that target other components of the mTOR pathway, promises to greatly aid research, while also potentially providing drugs of therapeutic value. Until recently, pharmacological inhibition of mTOR has been largely limited to rapamycins, but newer ATP-competitive inhibitors of mTOR kinase activity have been developed that block the phosphorylation of all known substrates of both mTORC1 and mTORC2 (ref. 73). These drugs show early promise as anticancer therapies, but as yet there are no data regarding their effects on longevity and other age-related diseases. S6K inhibitors, 4E-BP1 activators, mTORC1-specific inhibitors and mTOR-independent activators of autophagy are also attractive approaches for potentially modulating longevity and healthspan without incurring the side effects that are associated with rapamycins.

The search for a way to delay human ageing has proven long and elusive. Although still far from certain, there is reason to be optimistic that mTOR inhibitors may accomplish this goal. mTORC1 inhibition slows ageing in yeast and invertebrates, extends lifespan in mice, and has an impact on a diverse array of age-related diseases. Are we finally on the threshold of being able to fundamentally alter human ageing? Only time will tell, but if the pace and direction of recent progress are any indication, the next few chapters in the mTOR story should prove very interesting indeed. ■

Received 25 June; accepted 13 November 2012.

1. Laplante, M. & Sabatini, D. M. mTOR signaling in growth control and disease. *Cell* **149,** 274–293 (2012).
2. Stanfel, M. N., Shamieh, L. S., Kaeberlein, M. & Kennedy, B. K. The TOR pathway comes of age. *Biochim. Biophys. Acta* **1790,** 1067–1074 (2009).
3. Lamming, D. W. *et al.* Rapamycin-induced insulin resistance is mediated by mTORC2 loss and uncoupled from longevity. *Science* **335,** 1638–1643 (2012). **This study uses genetic models to confirm that reduced mTORC1 activity extends lifespan in mammals, and reports that the insulin resistance associated with chronic mTORC1 inhibition is caused by altered mTORC2 signalling and is not causally involved in lifespan extension.**
4. Fabrizio, P., Pozza, F., Pletcher, S. D., Gendron, C. M. & Longo, V. D. Regulation of longevity and stress resistance by Sch9 in yeast. *Science* **292,** 288–290 (2001).
5. Jia, K., Chen, D. & Riddle, D. L. The TOR pathway interacts with the insulin signaling pathway to regulate *C. elegans* larval development, metabolism and life span. *Development* **131,** 3897–3906 (2004).
6. Vellai, T. *et al.* Influence of TOR kinase on lifespan in *C. elegans*. *Nature* **426,** 620 (2003).
7. Kapahi, P. *et al.* Regulation of lifespan in *Drosophila* by modulation of genes in the TOR signaling pathway. *Curr. Biol.* **14,** 885–890 (2004).
8. Kaeberlein, M. *et al.* Regulation of yeast replicative life span by TOR and Sch9 in response to nutrients. *Science* **310,** 1193–1196 (2005).
9. Powers, R. W. III, Kaeberlein, M., Caldwell, S. D., Kennedy, B. K. & Fields, S. Extension of chronological life span in yeast by decreased TOR pathway signaling. *Genes Dev.* **20,** 174–184 (2006).
10. Medvedik, O., Lamming, D. W., Kim, K. D. & Sinclair, D. A. MSN2 and MSN4 link calorie restriction and TOR to sirtuin-mediated lifespan extension in *Saccharomyces cerevisiae*. *PLoS Biol.* **5,** e261 (2007).
11. Robida-Stubbs, S. *et al.* TOR signaling and rapamycin influence longevity by regulating SKN-1/Nrf and DAF-16/FoxO. *Cell Metab.* **15,** 713–724 (2012).
12. Bjedov, I. *et al.* Mechanisms of life span extension by rapamycin in the fruit fly *Drosophila melanogaster*. *Cell Metab.* **11,** 35–46 (2010).
13. Harrison, D. E. *et al.* Rapamycin fed late in life extends lifespan in genetically heterogeneous mice. *Nature* **460,** 392–395 (2009). **This study reports lifespan extension from mTORC1 inhibition in a mammal.**

14. Miller, R. A. *et al.* Rapamycin, but not resveratrol or simvastatin, extends life span of genetically heterogeneous mice. *J. Gerontol. A Biol. Sci. Med. Sci.* **66,** 191–201 (2011).
15. Anisimov, V. N. *et al.* Rapamycin increases lifespan and inhibits spontaneous tumorigenesis in inbred female mice. *Cell Cycle* **10,** 4230–4236 (2011).
16. Fontana, L., Partridge, L. & Longo, V. D. Extending healthy lifespan — from yeast to humans. *Science* **328,** 321–326 (2010).
17. Hansen, M. *et al.* Lifespan extension by conditions that inhibit translation in *Caenorhabditis elegans*. *Aging Cell* **6,** 95–110 (2007).
18. Zid, B. M. *et al.* 4E-BP extends lifespan upon dietary restriction by enhancing mitochondrial activity in *Drosophila*. *Cell* **139,** 149–160 (2009).
19. Kenyon, C. J. The genetics of ageing. *Nature* **464,** 504–512 (2010).
20. Takano, A. *et al.* Mammalian target of rapamycin pathway regulates insulin signaling via subcellular redistribution of insulin receptor substrate 1 and integrates nutritional signals and metabolic signals of insulin. *Mol. Cell. Biol.* **21,** 5050–5062 (2001).
21. Khatri, S., Yepiskoposyan, H., Gallo, C. A., Tandon, P. & Plas, D. R. FOXO3a regulates glycolysis via transcriptional control of tumor suppressor TSC1. *J. Biol. Chem.* **285,** 15960–15965 (2010).
22. Tettweiler, G., Miron, M., Jenkins, M., Sonenberg, N. & Lasko, P. F. Starvation and oxidative stress resistance in *Drosophila* are mediated through the eIF4E-binding protein, d4E-BP. *Genes Dev.* **19,** 1840–1843 (2005).
23. Guertin, D. A. *et al.* Ablation in mice of the mTORC components raptor, rictor, or mLST8 reveals that mTORC2 is required for signaling to Akt-FOXO and PKCα, but not S6K1. *Dev. Cell* **11,** 859–871 (2006).
24. Pan, K. Z. *et al.* Inhibition of mRNA translation extends lifespan in *Caenorhabditis elegans*. *Aging Cell* **6,** 111–119 (2007).
25. Apfeld, J., O'Connor, G., McDonagh, T., DiStefano, P. S. & Curtis, R. The AMP-activated protein kinase AAK-2 links energy levels and insulin-like signals to lifespan in *C. elegans*. *Genes Dev.* **18,** 3004–3009 (2004).
26. Onken, B. & Driscoll, M. Metformin induces a dietary restriction-like state and the oxidative stress response to extend *C. elegans* healthspan via AMPK, LKB1, and SKN-1. *PLoS ONE* **5,** e8758 (2010).
27. Anisimov, V. N. Metformin for aging and cancer prevention. *Aging (Albany NY)* **2,** 760–774 (2010).
28. Inoki, K., Zhu, T. & Guan, K. L. TSC2 mediates cellular energy response to control cell growth and survival. *Cell* **115,** 577–590 (2003).
29. Gwinn, D. M. *et al.* AMPK phosphorylation of raptor mediates a metabolic checkpoint. *Mol. Cell* **30,** 214–226 (2008).
30. Mair, W. *et al.* Lifespan extension induced by AMPK and calcineurin is mediated by CRTC-1 and CREB. *Nature* **470,** 404–408 (2011).
31. Leiser, S. F. & Kaeberlein, M. The hypoxia-inducible factor HIF-1 functions as both a positive and negative modulator of aging. *Biol. Chem.* **391,** 1131–1137 (2010).
32. Hudson, C. C. *et al.* Regulation of hypoxia-inducible factor 1α expression and function by the mammalian target of rapamycin. *Mol. Cell. Biol.* **22,** 7004–7014 (2002).
33. Mehta, R. *et al.* Proteasomal regulation of the hypoxic response modulates aging in *C. elegans*. *Science* **324,** 1196–1198 (2009).
34. Zhang, Y., Shao, Z., Zhai, Z., Shen, C. & Powell-Coffman, J. A. The HIF-1 hypoxia-inducible factor modulates lifespan in *C. elegans*. *PLoS ONE* **4,** e6348 (2009).
35. Leiser, S. F., Begun, A. & Kaeberlein, M. HIF-1 modulates longevity and healthspan in a temperature-dependent manner. *Aging Cell* **10,** 318–326 (2011).
36. Chen, D., Thomas, E. L. & Kapahi, P. HIF-1 modulates dietary restriction-mediated lifespan extension via IRE-1 in *Caenorhabditis elegans*. *PLoS Genet.* **5,** e1000486 (2009).
37. Selman, C. *et al.* Ribosomal protein S6 kinase 1 signaling regulates mammalian lifespan. *Science* **326,** 140–144 (2009).
This study shows that knockout of the mTORC1 substrate S6K1 extends lifespan in mice.
38. Colman, R. J. *et al.* Caloric restriction delays disease onset and mortality in rhesus monkeys. *Science* **325,** 201–204 (2009).
39. Fries, J. F. Aging, natural death, and the compression of morbidity. *N. Engl. J. Med.* **303,** 130–135 (1980).
40. Wilkinson, J. E. *et al.* Rapamycin slows aging in mice. *Aging Cell* **11,** 675–682 (2012).
41. Halloran, J. *et al.* Chronic inhibition of mammalian target of rapamycin by rapamycin modulates cognitive and non-cognitive components of behavior throughout lifespan in mice. *Neuroscience* **223,** 102–113 (2012).
42. Majumder, S. *et al.* Lifelong rapamycin administration ameliorates age-dependent cognitive deficits by reducing IL-1β and enhancing NMDA signaling. *Aging Cell* **11,** 326–335 (2012).
References 41 and 42 report that rapamycin improves cognitive function and protects against age-related cognitive decline in mice.
43. Kaeberlein, M. & Kennedy, B. K. Hot topics in aging research: protein translation and TOR signaling. *Aging Cell* **10,** 185–190 (2011).
44. Smith, E. D. *et al.* Quantitative evidence for conserved longevity pathways between divergent eukaryotic species. *Genome Res.* **18,** 564–570 (2008).
45. Kaeberlein, M. & Kennedy, B. K. Protein translation. *Aging Cell* **6,** 731–734 (2007).
46. Blagosklonny, M. V. Aging and immortality: quasi-programmed senescence and its pharmacologic inhibition. *Cell Cycle* **5,** 2087–2102 (2006).
47. Steffen, K. K. *et al.* Yeast lifespan extension by depletion of 60S ribosomal subunits is mediated by Gcn4. *Cell* **133,** 292–302 (2008).
48. Rogers, A. N. *et al.* Lifespan extension via eIF4G inhibition is mediated by post-transcriptional remodeling of stress response gene expression in *C. elegans*. *Cell Metab.* **14,** 55–66 (2011).
49. Thoreen, C. C. *et al.* A unifying model for mTORC1-mediated regulation of mRNA translation. *Nature* **485,** 109–113 (2012).
50. Mizushima, N., Levine, B., Cuervo, A. M. & Klionsky, D. J. Autophagy fights disease through cellular self-digestion. *Nature* **451,** 1069–1075 (2008).
51. Cuervo, A. M. Autophagy and aging: keeping that old broom working. *Trends Genet.* **24,** 604–612 (2008).
52. Hansen, M. *et al.* A role for autophagy in the extension of lifespan by dietary restriction in *C. elegans*. *PLoS Genet.* **4,** e24 (2008).
53. Alvers, A. L. *et al.* Autophagy is required for extension of yeast chronological life span by rapamycin. *Autophagy* **5,** 847–849 (2009).
54. Wei, M. *et al.* Life span extension by calorie restriction depends on Rim15 and transcription factors downstream of Ras/PKA, Tor, and Sch9. *PLoS Genet.* **4,** e13 (2008).
55. Bishop, N. A. & Guarente, L. Two neurons mediate diet-restriction-induced longevity in *C. elegans*. *Nature* **447,** 545–549 (2007).
56. Steinbaugh, M. J., Sun, L. Y., Bartke, A. & Miller, R. A. Activation of genes involved in xenobiotic metabolism is a shared signature of mouse models with extended lifespan. *Am. J. Physiol. Endocrinol. Metab.* **303,** E488–E495 (2012).
57. Cunningham, J. T. *et al.* mTOR controls mitochondrial oxidative function through a YY1–PGC-1α transcriptional complex. *Nature* **450,** 736–740 (2007).
58. Bonawitz, N. D., Chatenay-Lapointe, M., Pan, Y. & Shadel, G. S. Reduced TOR signaling extends chronological lifespan via increased respiration and upregulation of mitochondrial gene expression. *Cell Metab.* **5,** 265–277 (2007).
59. Pan, Y., Schroeder, E. A., Ocampo, A., Barrientos, A. & Shadel, G. S. Regulation of yeast chronological lifespan by TORC1 via adaptive mitochondrial ROS signaling. *Cell Metab.* **13,** 668–678 (2011).
60. Polak, P. *et al.* Adipose-specific knockout of raptor results in lean mice with enhanced mitochondrial respiration. *Cell Metab.* **8,** 399–410 (2008).
61. Chung, H. Y. *et al.* Molecular inflammation: underpinnings of aging and age-related diseases. *Ageing Res. Rev.* **8,** 18–30 (2009).
62. Morgan, T. E., Wong, A. M. & Finch, C. E. Anti-inflammatory mechanisms of dietary restriction in slowing aging processes. *Interdiscip. Top. Gerontol.* **35,** 83–97 (2007).
63. Liu, Y. Rapamycin and chronic kidney disease: beyond the inhibition of inflammation. *Kidney Int.* **69,** 1925–1927 (2006).
64. Nuhrenberg, T. G. *et al.* Rapamycin attenuates vascular wall inflammation and progenitor cell promoters after angioplasty. *FASEB J.* **19,** 246–248 (2005).
65. Chen, W. Q. *et al.* Oral rapamycin attenuates inflammation and enhances stability of atherosclerotic plaques in rabbits independent of serum lipid levels. *Br. J. Pharmacol.* **156,** 941–951 (2009).
66. Abdulrahman, B. A. *et al.* Autophagy stimulation by rapamycin suppresses lung inflammation and infection by *Burkholderia cenocepacia* in a model of cystic fibrosis. *Autophagy* **7,** 1359–1370 (2011).
67. Sharpless, N. E. & DePinho, R. A. How stem cells age and why this makes us grow old. *Nature Rev. Mol. Cell Biol.* **8,** 703–713 (2007).
68. Chen, C., Liu, Y. & Zheng, P. mTOR regulation and therapeutic rejuvenation of aging hematopoietic stem cells. *Sci. Signal.* **2,** ra75 (2009).
69. Yilmaz, O. H. *et al.* mTORC1 in the Paneth cell niche couples intestinal stem-cell function to calorie intake. *Nature* **486,** 490–495 (2012).
This study shows that rapamycin protects against loss of intestinal stem-cell function during ageing by preserving the stem-cell niche.
70. Cerletti, M., Jang, Y. C., Finley, L. W., Haigis, M. C. & Wagers, A. J. Short-term calorie restriction enhances skeletal muscle stem cell function. *Cell Stem Cell* **10,** 515–519 (2012).
71. Ramos, F. J. & Kaeberlein, M. A healthy diet for stem cells. *Nature* **486,** 477–478 (2012).
72. Singh-Manoux, A. *et al.* Timing of onset of cognitive decline: results from Whitehall II prospective cohort study. *Br. Med. J.* **344,** d7622 (2012).
73. Benjamin, D., Colombi, M., Moroni, C. & Hall, M. N. Rapamycin passes the torch: a new generation of mTOR inhibitors. *Nature Rev. Drug Discov.* **10,** 868–880 (2011).
74. Syntichaki, P., Troulinaki, K. & Tavernarakis, N. eIF4E function in somatic cells modulates ageing in *Caenorhabditis elegans*. *Nature* **445,** 922–926 (2007).
75. Curran, S. P. & Ruvkun, G. Lifespan regulation by evolutionarily conserved genes essential for viability. *PLoS Genet.* **3,** e56 (2007).
76. Chiocchetti, A. *et al.* Ribosomal proteins Rpl10 and Rps6 are potent regulators of yeast replicative life span. *Exp. Gerontol.* **42,** 275–286 (2007).
77. Malagelada, C., Jin, Z. H., Jackson-Lewis, V., Przedborski, S. & Greene, L. A. Rapamycin protects against neuron death in *in vitro* and *in vivo* models of Parkinson's disease. *J. Neurosci.* **30,** 1166–1175 (2010).
78. Tain, L. S. *et al.* Rapamycin activation of 4E-BP prevents parkinsonian dopaminergic neuron loss. *Nature Neurosci.* **12,** 1129–1135 (2009).
79. Majumder, S., Richardson, A., Strong, R. & Oddo, S. Inducing autophagy by rapamycin before, but not after, the formation of plaques and tangles ameliorates cognitive deficits. *PLoS ONE* **6,** e25416 (2011).
80. Spilman, P. *et al.* Inhibition of mTOR by rapamycin abolishes cognitive deficits and reduces amyloid-β levels in a mouse model of Alzheimer's disease. *PLoS ONE* **5,** e9979 (2010).
81. Wang, I. F. *et al.* Autophagy activators rescue and alleviate pathogenesis of a mouse model with proteinopathies of the TAR DNA-binding protein 43. *Proc. Natl Acad. Sci. USA* **109,** 15024–15029 (2012).
82. Garber, K. Targeting mTOR: something old, something new. *J. Natl Cancer Inst.* **101,** 288–290 (2009).
83. Shioi, T. *et al.* Rapamycin attenuates load-induced cardiac hypertrophy in mice. *Circulation* **107,** 1664–1670 (2003).
84. McMullen, J. R. *et al.* Inhibition of mTOR signaling with rapamycin regresses established cardiac hypertrophy induced by pressure overload. *Circulation* **109,** 3050–3055 (2004).

85. Ding, Y. *et al.* Haploinsufficiency of target of rapamycin attenuates cardiomyopathies in adult zebrafish. *Circ. Res.* **109,** 658–669 (2011).
86. Ramos, F. J. *et al.* Rapamycin reverses elevated mTORC1 signaling in lamin A/C-deficient mice, rescues cardiac and skeletal muscle function, and extends survival. *Sci. Transl. Med.* **4,** 144ra103 (2012).
87. Um, S. H. *et al.* Absence of S6K1 protects against age- and diet-induced obesity while enhancing insulin sensitivity. *Nature* **431,** 200–205 (2004).
88. Yang, S. B. *et al.* Rapamycin ameliorates age-dependent obesity associated with increased mTOR signaling in hypothalamic POMC neurons. *Neuron* **75,** 425–436 (2012).
89. Jagannath, C. & Bakhru, P. Rapamycin-induced enhancement of vaccine efficacy in mice. *Methods Mol. Biol.* **821,** 295–303 (2012).
90. Amiel, E. *et al.* Inhibition of mechanistic target of rapamycin promotes dendritic cell activation and enhances therapeutic autologous vaccination in mice. *J. Immunol.* **189,** 2151–2158 (2012).
91. Araki, K. *et al.* mTOR regulates memory CD8 T-cell differentiation. *Nature* **460,** 108–112 (2009).
92. Lieberthal, W. & Levine, J. S. Mammalian target of rapamycin and the kidney. II. Pathophysiology and therapeutic implications. *Am. J. Physiol. Renal Physiol.* **303,** F180–F191 (2012).
93. Kolosova, N. G. *et al.* Prevention of age-related macular degeneration-like retinopathy by rapamycin in rats. *Am. J. Pathol.* **181,** 472–477 (2012).
94. Yin, L., Ye, S., Chen, Z. & Zeng, Y. Rapamycin preconditioning attenuates transient focal cerebral ischemia/reperfusion injury in mice. *Int. J. Neurosci.* **122,** 748–756 (2012).
95. Tsai, P. T. *et al.* Autistic-like behaviour and cerebellar dysfunction in Purkinje cell *Tsc1* mutant mice. *Nature* **488,** 647–651(2012).
96. Ryther, R. C. & Wong, M. Mammalian target of rapamycin (mTOR) inhibition: potential for antiseizure, antiepileptogenic, and epileptostatic therapy. *Curr. Neurol. Neurosci. Rep.* **12,** 410–418 (2012).
97. Yamaki, K. & Yoshino, S. Preventive and therapeutic effects of rapamycin, a mammalian target of rapamycin inhibitor, on food allergy in mice. *Allergy* **67,** 1259–1270 (2012).
98. Pierdominici, M., Vacirca, D., Delunardo, F. & Ortona, E. mTOR signaling and metabolic regulation of T cells: new potential therapeutic targets in autoimmune diseases. *Curr. Pharm. Des.* **17,** 3888–3897 (2011).
99. Nussenblatt, R. B. *et al.* A randomized pilot study of systemic immunosuppression in the treatment of age-related macular degeneration with choroidal neovascularization. *Retina* **30,** 1579–1587 (2010).
100. Cao, K. *et al.* Rapamycin reverses cellular phenotypes and enhances mutant protein clearance in Hutchinson–Gilford progeria syndrome cells. *Sci. Transl. Med.* **3,** 89ra58 (2011).

Acknowledgements Studies related to this topic in the Kaeberlein laboratory have been supported by NIH grant R01AG039390 and R01AG031108 to M.K. S.C.J. is supported by NIH grant T32AG000057.

Author Information Reprints and permission information is available at www.nature.com/reprints. The authors declare no competing financial interests. Readers are welcome to comment on the online version of this article at go.nature.com/e8iwng. Correspondence should be addressed to M.K. (kaeber@u.washington.edu).

doi:10.1038/nature11862

Metabolism of inflammation limited by AMPK and pseudo-starvation

Luke A. J. O'Neill[1] & D. Grahame Hardie[2]

Metabolic changes in cells that participate in inflammation, such as activated macrophages and T-helper 17 cells, include a shift towards enhanced glucose uptake, glycolysis and increased activity of the pentose phosphate pathway. Opposing roles in these changes for hypoxia-inducible factor 1α and AMP-activated protein kinase have been proposed. By contrast, anti-inflammatory cells, such as M2 macrophages, regulatory T cells and quiescent memory T cells, have lower glycolytic rates and higher levels of oxidative metabolism. Some anti-inflammatory agents might act by inducing, through activation of AMP-activated protein kinase, a state akin to pseudo-starvation. Altered metabolism may thus participate in the signal-directed programs that promote or inhibit inflammation.

To many biologists, metabolism is viewed simply as a means to generate a store of energy by catabolism, and to generate macromolecules for cell maintenance and growth through anabolic pathways. The elucidation of metabolic pathways in the twentieth century gave insight into disorders in which there are obvious dysfunctions in metabolism, such as diabetes and atherosclerosis. However, in the twenty-first century, alterations in metabolic regulation are now seen to be just as important in other diseases, such as cancer and inflammatory conditions. As long ago as 1923, the German biochemist Otto Warburg proposed that metabolic dysregulation was a feature of tumour cells[1]. Normal cells metabolize glucose through glycolysis, and under normoxic conditions, the pyruvate that is generated is then further oxidized in mitochondria. When oxygen becomes limiting, however, mitochondrial oxidative metabolism is restricted, and the pyruvate is converted to lactate instead. Warburg discovered that this latter process predominates in tumour cells, even when oxygen is plentiful, a process that became known as the 'Warburg effect' or aerobic glycolysis.

Studies on the metabolism of activated macrophages began in the 1960s. Monocytes from peritoneal exudates were shown to depend mainly on glycolysis as a source of metabolic energy[2], whereas *in vitro* culture of monocytes caused a significant increase in glycolysis[3]. In 1982, a study[4] showed that in an *in vivo* model (graft versus host) of immune activation, lymphocytes exhibited increased glycolysis — which was shown to be important for proliferation — together with a marked increase in glutamine use[4]. Activated macrophages have been shown to have a high hexokinase activity, the first enzyme involved in glycolysis and in the pentose phosphate pathway. Glycolysis and glutamine metabolism are also markedly increased during phagocytosis[5].

More recent studies have confirmed and extended these earlier findings. A shift towards aerobic glycolysis occurs in macrophages and dendritic cells activated by the bacterial product lipopolysaccharide, acting through Toll-like receptor 4 (TLR4)[6], in M1 inflammatory macrophages[7], and in T-helper (T_H) 17 lymphocytes that produce the pro-inflammatory cytokine interleukin-17 (IL-17)[8,9]. On the other hand, cells that limit inflammation, such as regulatory T cells[8,9], M2 anti-inflammatory macrophages[7] and quiescent memory T cells that carry the CD8 antigen[10], exhibit oxidative metabolism with more limited rates of glycolysis. Mechanistic insight into these changes has lagged behind tumour cells, but they are likely to share several features, such as the involvement of the transcription factor hypoxia-inducible factor 1α (HIF-1α) and AMP-activated protein kinase (AMPK). Here we discuss these recent developments in the case of inflammation, compare them with findings obtained with tumour cells and speculate on what this new information might mean for our understanding of the pathogenesis of inflammatory diseases, and for possible new directions for therapeutics.

Inflammation and metabolism

Inflammation is triggered in response to tissue injury or infection, and mounting an inflammatory response is an energy-intensive process. In the presence of pathogens or the products of inflamed tissues that provoke inflammation, macrophages rapidly switch from a resting state to a highly active state, exhibiting a pronounced increase in production of host defence factors, enhanced phagocytosis and antigen presentation. It is perhaps no surprise that such highly active cells undergo metabolic changes similar to those seen in tumour cells. The similarity between the metabolism that occurs in tumour cells and activated T cells has already been pointed out, with particular roles for the phosphatidylinositol-3-OH kinase (PI(3)K)–Akt and the mechanistic target of rapamycin (mTOR) pathways, as well as the transcription factor c-myc[11]. One difference between metabolic pathways in tumour cells and inflammatory cells (Fig. 1) is that the changes in tumour cells are driven largely by mutations, whereas those in inflammatory cells are driven by extracellular signals.

Macrophage and dendritic-cell activation

Resting dendritic cells predominantly use oxidative phosphorylation to generate ATP, but on activation of TLRs — which are the primary sensors of microbial products that trigger inflammation by pathogen-associated molecular patterns (PAMPs) — they undergo a switch to glycolysis[6] (Fig. 2). Such activated dendritic cells have elevated expression of the glucose transporter GLUT1 and increased lactate production, with a decrease in oxygen consumption by mitochondria — effects that are dependent on both PI(3)K and Akt1 (ref. 6). As discussed later, these effects are opposed by AMPK activation. Also of note is the effect of the immunomodulatory cytokine IL-10, which limits the switch to glycolysis and blocks TLR action in dendritic cells, again suggesting that a switch to mitochondrial metabolism has a net anti-inflammatory effect[6]. Depression of oxidative metabolism by TLR4 involves upregulation of the inducible form of nitric oxide synthase (iNOS)[12]; NO competes with

---

[1] School of Biochemistry and Immunology, Trinity Biomedical Sciences Institute, Trinity College Dublin, Ireland. [2] Division of Cell Signalling and Immunology, College of Life Sciences, University of Dundee, Dundee DD1 5EH, Scotland, UK.

oxygen to inhibit the terminal electron acceptor of the respiratory chain, cytochrome *c* oxidase. By preventing reoxidation of NADH, this limits flux through the tricarboxylic acid (TCA) cycle, leading to an increase in pyruvate that is diverted to generate lactate, an alternative route for reoxidizing the NADH produced by glycolysis. The increase in glycolysis therefore occurs in concert with the attenuation of oxidative phosphorylation, and serves as an alternative pathway to maintain ATP levels. ATP is required not only by many ATP-dependent enzymes, but also to maintain mitochondrial membrane potential through reversal of adenine nucleotide translocase and the F_OF_1-ATP synthase when NO inhibits the respiratory chain[13]. Without this, a drop in the mitochondrial membrane potential would cause recruitment of the apoptosis regulator BAX to the mitochondria, release of cytochrome *c* into the cytosol, and consequent activation of caspases and apoptotic cell death. Enhanced ATP levels could, therefore, prolong the lifespan of dendritic cells and promote adaptive immunity.

Apart from the attenuation in oxidative phosphorylation in macrophages in response to TLR stimulation, the production of reactive oxygen species (ROS) from mitochondria increases. Mitochondria are recruited to phagosomes that contain bacteria, and the ROS that they produce augment the bactericidal activity of macrophages[14]. This provides an alternative source of ROS to the NADPH oxidase system, with current evidence indicating that mitochondrial ROS are required for macrophage bactericidal activity[14,15]. The TLR-signalling protein TRAF6 is recruited to mitochondria on activation of TLR2 or TLR4, but not TLR3 or TLR9, and engages with the adaptor protein ECSIT on the mitochondrial surface. Both TRAF6 and ECSIT are required for the generation of ROS[14]. The different effects of specific TLRs on this process can be rationalized. The TLR1–TLR2 dimer senses bacterial lipopeptides and TLR4 senses lipopolysaccharide, whereas TLR3 and TLR9 sense double-stranded RNA and CpG-rich DNA, respectively, both of which are common in viruses. TLR1–TLR2 and TLR4 also signal from the plasma membrane, whereas TLR3 and TLR9 signal from endosomes. In TLR1–TLR2 and TLR4 signalling, induction of ROS is required to kill bacteria in phagosomes, whereas for TLR3 and TLR9 this is not required for antiviral effects.

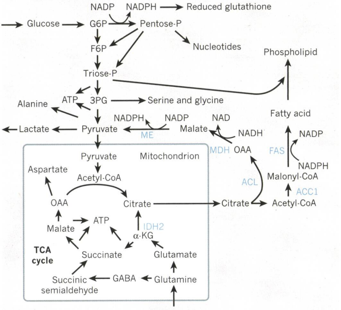


Figure 1 | Central metabolic pathways involved in the Warburg effect in tumour cells and inflammatory cells. Glucose is taken up by the cell and converted to pyruvate by glycolysis; the pyruvate then enters the energy-generating TCA cycle. The pentose phosphate pathway generates reduced glutathione and nucleotides. Citrate produced by the TCA cycle is converted into fatty acids and phospholipids. Glutamine is taken up by the cell and converted, by glutaminolysis, to glutamate or GABA (γ-aminobutyric acid) — produced by the GABA shunt — for entry into the TCA cycle. Enzymes involved in the pathways are shown in blue. ACC1, acetyl-coenzyme A (CoA) carboxylase 1; ACL, ATP citrate lyase; FAS, fatty-acid synthase; F6P, fructose-6-phosphate; G6P, glucose-6-phosphate; IDH2, isocitrate dehydrogenase 2; α-KG, α-ketoglutarate; MDH, malate dehydrogenase; ME, malic enzyme; OAA, oxaloacetate; P, phosphate; 3PG, 3-phosphoglycerate.

Mechanisms of metabolic changes in macrophages

If increased glycolysis is required because of a decrease in oxidative phosphorylation in dendritic cells and macrophages, how is it switched on? Here the similarity to tumour cells becomes clear. Macrophages activated with lipopolysaccharide and interferon-γ (IFN-γ) acquire an innate or inflammatory phenotype, and are often termed M1 macrophages[7]. They have a high microbicidal activity, as well as enhanced production of pro-inflammatory cytokines and ROS. There is also a marked switch from expression of the liver isoform of 6-phosphofructo-2-kinase (encoded by *PFKFB1*) to the PFKFB3 isoform, the type also commonly found in tumour cells[7]. This leads to accumulation of fructose-2,6-bisphosphate and, thus, increased glycolytic flux. The opposite is true when macrophages are activated by IL-4 and IL-13, which promote the alternative phenotype, often termed M2 macrophages. These macrophages are associated with tissue repair and humoral immunity, as well as anti-inflammatory cytokine production and reduced expression of major histocompatibility complex (MHC) class II and therefore antigen presentation. They exhibit enhanced oxidative phosphorylation and much lower rates of glycolysis[7], and have no detectable PFKFB3, expressing PFKFB1 instead. M2 macrophages still produce cytokines, but these are largely anti-inflammatory in action. Why oxidative phosphorylation is associated with production of anti-inflammatory cytokines, whereas glycolysis is associated with production of pro-inflammatory cytokines, remains unclear. However, peroxisome-proliferator-activated receptor-γ co-activator-1β (PGC-1β) has a key role in the mechanism[16]. PGC-1β is a transcriptional co-activator that promotes oxidative metabolism, notably by upregulating expression of genes involved in fatty-acid oxidation[17]. In macrophages activated with IL-4 to generate the M2 phenotype, there is a profound increase in the entire program for fatty-acid metabolism, including uptake and oxidation of fatty acids and mitochondrial biogenesis (Fig. 3a). These changes are dependent on activation by IL-4 of transcription factor STAT6, which in turn induces PGC-1β (ref. 16). Over-expression of PGC-1β promotes the M2 phenotype and attenuates macrophage-mediated inflammation *in vivo*. Conversely, knockdown of the gene that encodes PGC-1β limits the shift to oxidative phosphorylation and the function of M2 macrophages, and greatly boosts the activating effects of lipopolysaccharide[16]. STAT6 and PGC-1β have been shown to interact directly at the promoters of genes expressed specifically in M2 macrophages, such as ARG1 (ref. 16). A mechanism has therefore been defined whereby cytokines such as IL-4, acting through STAT6, promote PGC-1β function, leading to increased fatty-acid oxidation, mitochondrial biogenesis and oxidative phosphorylation (Fig. 3a).

In another study, IFN-γ, acting through STAT1, was shown to induce ROS through the nuclear receptor oestrogen-related receptor-α (ERR-α), which in turn was dependent on PGC-1β. ERR-α deficient mice were shown to be more susceptible to infection with *Listeria monocytogenes*[15]. This study further emphasizes the importance of mitochondrial ROS in killing bacteria.

Function of sirtuins in macrophages

Sirtuins are a family of NAD$^+$-dependent lysine deacetylases that act on a wide range of protein substrates. They are thought to act as NAD$^+$ sensors that connect nutrition and metabolism to chromatin structure[18]. Sirtuin activation has been shown to be anti-inflammatory, and knockdown experiments in macrophages have shown that two sirtuins, Sirt1 and Sirt6, promote a switch from glycolysis to fatty-acid oxidation[19], which requires the enzyme nicotinamide phosphoribosyltransferase

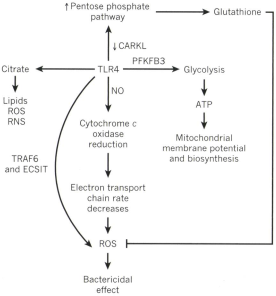

↑ Pentose phosphate pathway
Glutathione
↓ CARKL
Citrate
TLR4
PFKFB3
Glycolysis
Lipids
ROS
RNS
NO
ATP
Cytochrome *c* oxidase reduction
Mitochondrial membrane potential and biosynthesis
TRAF6 and ECSIT
Electron transport chain rate decreases
ROS
Bactericidal effect

Figure 2 | Metabolic changes in TLR4-activated M1 macrophages. The lipopolysaccharide receptor TLR4 promotes a major shift in the metabolism of macrophages and dendritic cells, which is a determining factor in their activation. TLR4 induces the enzymes involved in glycolysis, including the key controlling enzyme PFKFB3; the ATP produced is required for biosynthesis and to maintain mitochondrial membrane potential. TLR4 activation also results in oxidative metabolism attenuation in mitochondria, mediated in part by nitric oxide (NO), causing the displacement of oxygen from cytochrome *c* oxidase. This leads to a limitation in the rate of transfer in the electron transport chain, generating reactive oxygen species (ROS), which have a bactericidal effect. An increase in ROS is promoted by the TLR4 signals TRAF6 and ECSIT. The activity of the pentose phosphate pathway increases owing to a decrease in the enzyme carbohydrate kinase-like protein (CARKL). The pentose phosphate pathway generates NADPH, which is needed for glutathione function — a key antioxidant that limits ROS. TLR4 also generates citrate, which is used in lipid biosynthesis, as well as to generate further ROS and reactive nitrogen species (RNS).

(NAMPT) — involved in the synthesis of NAD$^+$. Sirtuins are known to be involved in the effects of dietary restriction, and may partly explain why this manipulation is beneficial (because it would limit inflammation). Sirt1 deacetylates and inactivates the p65 component of the nuclear factor-κB (NF-κB) pathway[20], thus limiting the expression of NF-κB-dependent genes. Sirt1 also deacetylates and activates PGC-1β[21] (Fig. 3a), thus promoting a shift towards fatty-acid oxidation[22]. Activation of TLR4 will actually induce NAMPT late in the time course of macrophage activation, and thereby cause a negative feedback effect on macrophage activation[23]. Similarly, the anti-inflammatory effects of dietary restriction, or pharmacological activation of sirtuins, might occur through the reprogramming of macrophages to an anti-inflammatory state. Sirtuins were originally identified as having a role in prolonging lifespan, but the beneficial effects of sirtuin activation in mammals may owe more to their anti-inflammatory effects.

The role of carbohydrate kinase-like protein

In a screen for new regulators of macrophage activation, carbohydrate kinase-like protein (CARKL) was found to be downregulated in macrophages[24]. When CARKL is overexpressed, the induction of pro-inflammatory cytokines by lipopolysaccharide is repressed, whereas knocking down its gene had the opposite effect. This indicates that CARKL is a key driver of the M2 phenotype, with lipopolysaccharide suppressing its expression to promote the M1 phenotype. Intriguingly, CARKL is a sedoheptulose kinase that catalyses the formation of sedoheptulose 7-phosphate (S7P) (an intermediate in the pentose phosphate pathway), which seems to limit flux into the pentose phosphate pathway. This leads to a reduction in the formation of NADPH and hence reduced glutathione, perhaps leading to a pro-oxidant state that limits M1 macrophage function. When CARKL levels are low, as is triggered by TLR4, a more antioxidant environment prevails in the macrophage. IL-4 induces CARKL expression (Fig. 3a). The requirement for this antioxidant state in M1 macrophages may involve NF-κB, a key transcription factor for pro-inflammatory genes such as tumour necrosis factor-α (TNF-α), because NF-κB subunits such as RelA must remain in a reduced state to function properly[25].

This study on CARKL is important because it integrates several concepts. Firstly, a key feature of M1 macrophages may be their ability to produce ROS, which are needed to kill phagocytosed bacteria, from mitochondria. However, these ROS must not cause intracellular damage. Glycolysis can provide ATP to maintain mitochondrial membrane potential in the absence of full function of the respiratory chain, and to drive reactions involved in the biosynthesis of pro-inflammatory proteins. Increased glucose uptake also promotes the pentose phosphate pathway, which is needed, in part, to generate the NADPH required for maintenance of reduced glutathione, limiting the cellular damage that ROS might otherwise cause. Activated macrophages have increased expression of glucose-6-phosphate 1-dehydrogenase and 6-phosphogluconate dehydrogenase, two key enzymes in the pentose phosphate pathway. In addition, glutamine metabolism can generate NADPH, and a key enzyme in this process is glutamate dehydrogenase, which is also elevated in lipopolysaccharide-activated macrophages[26].

The roles of M1 and M2 macrophages

ROS production is a key functional feature of the acute inflammatory, antibacterial M1 macrophage, and the metabolic alterations that occur are integral to this process. The anti-inflammatory, tissue-repair-promoting M2 macrophage, however, has a quite different metabolism. It is likely that the metabolic switching is needed for the distinct functions of these macrophages. M1 macrophages need to have a rapid, vigorous response to bacteria, involving the generation of ROS and pro-inflammatory cytokines such as IL-1β and TNF-α, which are needed to deal with infection and danger from damaged tissue. The increase in ROS occurs partly because of attenuation in the respiratory chain, whereas increased glycolytic flux is needed to maintain ATP levels for biosynthesis, as well as to maintain mitochondrial membrane potential and prevent apoptosis. Equally, increased glucose uptake and glycolysis are needed to allow the induction of expression of inflammatory genes (many of which require HIF-1α, as discussed later), and the production of NADPH by the pentose phosphate pathway to limit the damage caused by ROS. By contrast, M2 macrophages are more involved in long-term tissue repair, and provide defence against extracellular parasites. STAT6 controls the genetic program for long-term macrophage activation. Oxidative metabolism could be better suited to these longer-term roles of M2 macrophages. This oxidative metabolism requires the STAT6 target gene encoding PGC-1β, whereas the M1 phenotype requires HIF-1α and glycolysis. Although M2 macrophages are highly active and can proliferate, they rely mainly on oxidative metabolism for their energetic and biosynthetic demands.

A role for citrate in macrophage activation

A final aspect of macrophage metabolism and inflammation concerns the role of citrate. Altered metabolism of citrate is an important feature of the metabolic switch that occurs in proliferating tumour cells, when the TCA cycle changes from being a purely catabolic pathway generating ATP to become, at least in part, an anabolic pathway. Citrate is withdrawn from the cycle (Fig. 1) to provide, through the reaction catalysed by ATP citrate lyase, a source of cytoplasmic acetyl-coenzyme A (CoA) for the synthesis of large amounts of fatty acids that are required for the synthesis of new membrane lipids — including phospholipids, which are the source of arachidonic acid for prostaglandin production. Another product of citrate metabolism is oxaloacetate, which produces NADPH through cytosolic malate dehydrogenase and malic enzyme (Fig. 1). NADPH is required to generate ROS through NADPH oxidase, and NO through NO synthase. Interestingly, knockdown of the mitochondrial citrate carrier gene limits the export of citrate and blocks the production of prostaglandins, ROS and NO in response to lipopolysaccharide[27]. Citrate derived from mitochondria is therefore required for certain responses to lipopolysaccharide.

T_H17 lymphocytes versus regulatory T cells

There are intriguing similarities between the metabolic profile of T lymphocytes involved in the inflammatory process, such as T_H17 lymphocytes, and those exhibited by inflammatory macrophages and dendritic cells. Equally, anti-inflammatory T lymphocytes such as regulatory T (T_{reg}) cells have a metabolism that is characterized by mitochondrial oxidative metabolism, analogous to M2 macrophages. We can postulate, therefore, that a shift towards high glycolysis is a hallmark of inflammatory cells, whereas oxidative phosphorylation is a hallmark of anti-inflammatory cells. Increased glycolysis in lymphocytes grown in culture was shown as early as 1966 (ref. 28). This has also been shown in cells that carry the CD4 antigen, in which CD28 signalling increases glycolysis through activation of PI(3)K and Akt[29]. Glucose uptake and glycolysis are increased in T_H17 cells (as well as T_H2 and T_H1 cells) compared with T_{reg} cells[30], which in turn have increased membrane potential and oxidize lipids at a higher rate than other subsets of cells that carry the CD4 antigen. There is, however, likely to be a gradation towards either glycolysis or oxidative metabolism because T_H17 cells still exhibit some lipid oxidation.

Shi and colleagues have reported that in conditions favouring T_H17-cell development from naive T cells, glycolysis is strongly upregulated and glycolytic enzymes are induced; however, this did not occur under conditions in which T_{reg} cells were induced[8]. Blocking glycolysis with 2-deoxyglucose led to a switch in lineage, with T_H17 cells becoming T_{reg} cells; this striking result suggests that the plasticity of T cells, a key theme in immunology, might be controlled by metabolic cues and activated by polarizing signals such as IL-17 or transforming growth factor-β (TGF-β). HIF-1α expression was also induced in T_H17 cells, and this required mTOR complex-1 (mTORC1), with rapamycin limiting these changes and promoting development of T_{reg} cells. A link between mTORC1 and HIF-1α is that mTORC1 enhances translation of the messenger RNA encoding HIF-1α, which contains 5′-terminal oligopyrimidine (5′-TOP) sequences[31]; translation of mRNAs containing such sequences is known to be upregulated following mTORC1 activation. The HIF-1α-dependent glycolytic pathway therefore provides a stimulus for the differentiation of T_H17 cells, and limits T_{reg} cell development[8].

The PI(3)K–mTOR pathway has been studied extensively in T-cell function. It has an essential role in promoting glucose use in effector T-cell development, including T_H17 cells, and at the same time inhibits T_{reg} cell generation[32,33]. mTORC1 acts downstream of Akt, while mTORC2 promotes Akt activation. T cells with a T-cell-specific mTOR knockout do not differentiate into effector T cells, including T_H17 cells, and default to the T_{reg} phenotype[34], which is consistent with results showing that rapamycin promotes fatty-acid oxidation and generates T_{reg} cells[35]. mTORC1 and mTORC2 also have specific effects on effector T-cell subsets. T cells deficient in the GTP-binding protein Rheb, which lack mTORC1 function, can only differentiate into T_H2 cells, whereas mTORC2-deficient T cells differentiate into T_H1 and T_H17, but not T_H2 cells[36]. The molecular basis of these differences and the role of metabolism in the effects of mTORC1 and mTORC2 deficiency are not known.

In a study similar to that of Shi *et al.*[8], Dang and colleagues demonstrated that HIF-1α enhances T_H17 cell development through the direct transcriptional activation of the key T_H17 transcription factor retinoic-acid-related orphan receptor-γt (RORγt), which then forms a complex with p300 on the IL-17 promoter[9] (Fig. 3b). One role of HIF-1α is therefore to promote the function of RORγt. HIF-1α also binds to the T_{reg} transcription factor Foxp3 and targets it for proteosomal degradation through recruitment of the E3 ligase VHL, which occurs under normoxic conditions. Crucially, therefore, HIF-1α is a molecular switch in T-cell differentiation, because it directly promotes RORγt, but limits Foxp3. The mechanism of HIF-1α induction has been shown to involve activation of STAT3 by IL-6, with STAT3 then binding to the HIF-1α promoter and increasing its transcription[7].

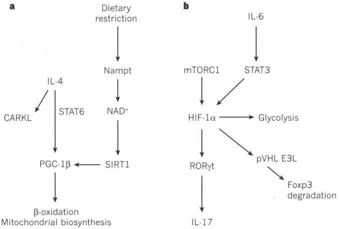


Figure 3 | Metabolic regulation of M2 macrophages and T_H17 cells. **a**, In M2 macrophages, the cytokine IL-4 acts through STAT6 to induce PGC-1β, a key transcription factor for β-oxidation in mitochondria and mitochondrial biogenesis. M2 macrophages are also activated by dietary restriction, through promotion of NAMPT which generates a key cofactor for SIRT1, NAD$^+$. This enzyme deactylates and activates PGC-1β. The M2 macrophage, therefore, generates mainly ATP from oxidative metabolism in mitochondria, which can use multiple carbon sources and produces less ROS compared with M1 macrophages. This reflects the role of M2 macrophages in longer term processes, such as tissue repair and antiparasitic effects. CARKL is a key driver of the M2 phenotype, and its expression is induced by IL-4. **b**, In T_H17 cell differentiation, IL-6 activates STAT3, which induces HIF-1α. mTORC1 is also involved in this process, as it stabilizes HIF-1α-encoding mRNA. HIF-1α induces RORγt, which is the key transcription factor for IL-17. HIF-1α also promotes degradation of the T_{reg} transcription factor Foxp3 through the von Hippel-Lindau E3 ligase (pVHL E3L), strengthening the phenotype of the T_H17 cell. It may also promote glycolysis, but there is currently no evidence for this.

Importantly, HIF-1α is not required to promote the glycolysis that occurs in initial T-cell activation, but influences instead the proportion of T_H17 cells relative to T_{reg} cells[8,9]. The transcription factor c-myc has been shown to be more important for glycolysis and glutamine metabolism in the initial phase of T-cell activation[11]. It is not clear why there are temporal differences in the regulation of glycolysis by these two important transcription factors.

From these studies, along with those in macrophages, we can conclude that HIF-1α promotes glycolysis and the pentose phosphate pathway through induction of the appropriate enzymes and transporters, and the function of this is perhaps to generate sufficient ATP in the face of decreased mitochondrial metabolism while promoting biosynthesis and antioxidant defence. This could be why HIF-1α plays a part in the later stages of T-cell activation, as it is activated as a result of decreased mitochondrial metabolism. In the case of macrophages, HIF-1α is also required to induce a set of pro-inflammatory genes. In the T cell, HIF-1α promotes the action of RORγt on the IL-17 promoter, but limits Foxp3 and prevents T_{reg} development. HIF-1α, therefore, has a pivotal role in both cases. However, similar to macrophages, the question arises why the more anti-inflammatory cell has a predominance of mitochondrial oxidative metabolism. Again the answer may lie partly in timing: inflammation must be activated rapidly to counter infection or deal with injury, so T_H17 cells emerge rapidly with a high glycolytic rate and the capacity to release the important pro-inflammatory cytokine IL-17. T_{reg} cells, however, will also emerge, become activated and proliferate, limiting inflammation

Table 1 | Metabolic characteristics of tumour and inflammatory cells

	Tumour cell	M1 macrophage	$T_{H}17$ cell	M2 macrophage	T_{reg} cell
Hallmark metabolic processes	Glycolysis and pentose phosphase pathway	Glycolysis and pentose phosphate pathway	Glycolysis	Oxidative phosphorylation	Oxidative phosphorylation
Key regulatory enzymes	PFKFB3 and mTORC1	PFKFB3 and mTORC1	mTORC1	PFKFB1, AMPK and sirtuins	NA
Transcription factors	HIF-1α	HIF-1α	STAT3, HIF-1α and RORγt	STAT6 and PGC-1β	Foxp3
CARKL status	Not described	Low	Not described	High	Not described
Use of citrate	Lipid biosynthesis, ROS and RNS	Lipid biosynthesis, ROS and RNS	NA	NA	NA
Overall phenotype	Tumour	Pro-inflammatory and bactericidal	Pro-inflammatory	Anti-inflammatory, tissue repair and antiparasitic	Anti-inflammatory and tissue repair

NA, not applicable; RNS, reactive nitrogen species; ROS, reactive oxygen species.

once the danger has passed, as well as promoting repair and recovery. This might require the sustained production of ATP from multiple carbon sources through oxidative phosphorylation, while limiting ROS production.

An important role for HIF-1α in inflammation can also be inferred from studies in myeloid cells. HIF-1α-deficient macrophages are attenuated in their capacity to kill bacteria, and HIF-1α is important for the production of antimicrobial peptides and granule proteases[37,38]. The transcription factor has also been shown to be crucial for controlling the inflammatory response of macrophages in skin and joints, and promoting their phagocytic activity[39]. HIF-1α-deficient mice are resistant to septic shock induced by lipopolysaccharide, and macrophages from HIF-1α-deficient mice make fewer pro-inflammatory cytokines[37] than those of wild-type mice. This is also observed in macrophages from HIF-2α-deficient mice suggesting that both isoforms of HIF are required for lipopolysaccharide responses *in vivo*, possibly acting in different cell types[40]. The overall role of HIF-1α can therefore be summarized as being required not only for glycolysis, but also in the induction of key pro-inflammatory proteins.

CD8⁺ memory T cells

Another example of the importance of mitochondrial respiration is in T cells that carry the CD8 antigen. The establishment of a memory population of T cells is a key feature of adaptive immunity. TRAF6 is known to be required for the transition from effector to memory T cells, with TRAF6-deficient T cells being unable to switch from glycolysis to fatty-acid oxidation[41], a result that suggests that this metabolic shift is required for the formation of memory T cells. In addition, rapamycin promotes the generation of memory T cells in mice that have been infected with a virus, again probably by promoting mitochondrial metabolism[42]. Van der Windt and colleagues studied mitochondrial metabolism in effector T cells that carry the CD8 antigen and compared this to memory cells[10]. They found that the memory population had substantial spare mitochondrial capacity, which was accounted for by the cytokine IL-15, a key to memory T-cell function. IL-15 caused mitochondrial biogenesis and expression of carnitine palmitoyl transferase-1 (CPT1), a regulatory enzyme involved in mitochondrial fatty-acid oxidation. The authors propose that mitochondrial respiratory capacity regulates T-cell survival after infection. Effector T cells, such as those that carry the CD8 antigen or $T_{H}17$ cells, use mainly glycolysis to maintain growth and function. However, the attenuated mitochondrial metabolism in these cells causes them to become unstable because they are unable to generate energy from diverse substrates through oxidative phosphorylation, and thus cannot maintain viability when infection clears. However, a small number of these T cells survive because IL-15, or similar cytokines such as IL-7, promote a switch to mitochondrial biogenesis, fatty-acid oxidation and survival, with these cells then persisting as memory T cells. The cells enter a state of quiescence using mitochondrial metabolism to generate ATP until a similar infection appears, when they rapidly undergo differentiation into effector T cells to deal with the infection. The main metabolic features of tumour cells, M1 and M2 macrophages, $T_{H}17$ and T_{reg} cells are shown in Table 1.

AMPK as a key signalling pathway

Insight into the role of metabolism in inflammation is supported by recent findings concerning the role of AMPK in inflammation and immunity; these findings also provide the enticing prospect of new therapeutic approaches for inflammatory diseases. AMPK is a highly conserved protein kinase that exists in essentially all eukaryotic cells in the form of heterotrimeric complexes comprising catalytic α-subunits and regulatory β- and γ-subunits. The α-subunits contain a conserved threonine residue (Thr 172 in rat α2) for which phosphorylation by upstream kinases causes more than 100-fold activation. The primary upstream kinase that phosphorylates Thr 172 is the protein kinase LKB1 (refs 43–45). This was an exciting finding because LKB1 was previously identified as a tumour suppressor that was mutated in the inherited cancer-susceptibility disorder Peutz–Jeghers syndrome, and in many sporadic cancers[46]. LKB1 has a high basal activity and, thus, phosphorylates Thr 172 on AMPK constitutively[47]. However, binding of adenine nucleotides to the regulatory γ-subunit of AMPK causes conformational changes that modulate the phosphorylation state of Thr 172, thereby adjusting the kinase activity of AMPK according to the energy status of the cell. Binding of ATP (which signifies adequate energy status) reduces net Thr 172 phosphorylation, whereas binding of ADP or AMP (which signifies impaired energy status) enhances net Thr 172 phosphorylation, increasing the kinase activity[48–50].

Evolutionary conservation of the role of AMPK

Genes encoding the three subunits of the AMPK heterotrimer can be easily recognized in almost all eukaryotic genomes. The orthologue in the yeast *Saccharomyces cerevisiae*, the SNF1 complex, is activated by glucose starvation and is required for growth at low glucose concentrations[51,52]. Interestingly, yeast exhibits a phenomenon that is closely related to the Warburg effect, although it generates ethanol rather than lactate. When grown at high glucose concentrations, it uses glycolysis to generate ethanol (through fermentation) to make ATP, and the expression of enzymes of oxidative metabolism is repressed[53]. Conversely, when yeast runs short of glucose, growth slows and it switches from fermentation to the more energy-efficient oxidative metabolism. This adaptation involves upregulation of genes that are required for oxidative metabolism and downregulation of genes that are required for glycolysis, and a functional SNF1 complex is required for most of these changes[54]. Thus, it can be argued that the SNF1 complex exerts an anti-Warburg effect, implying that the mammalian orthologue might do the same. The overall effects of AMPK activation on mammalian metabolism are, in general, compatible with the idea that it is a tumour suppressor that promotes the oxidative metabolism that is typical of quiescent or anti-inflammatory cells, rather than the glycolytic metabolism typical of tumour cells or inflammatory cells. Once switched on, AMPK attempts to restore energy homeostasis by activating catabolic pathways that generate ATP efficiently, while switching off energy-consuming processes such as biosynthesis and cell-cycle progression[55,56] (Fig. 4). A parallel can therefore be drawn between starving yeast cells and M2 macrophages, T_{reg} cells and memory T cells that carry the CD8 antigen, all of which are in a state characterized by oxidative metabolism, perhaps mediated partly by AMPK activation. If this is correct, AMPK-activating drugs will be anti-inflammatory, inducing what could be called a state of pseudo-starvation.

Catabolic processes that are switched on by AMPK include expression of genes involved in mitochondrial oxidative metabolism, as well as mitochondrial biogenesis itself — an effect mediated by covalent modification of PGC-1α, which is closely related to PGC-1β (ref. 57). AMPK triggers both phosphorylation[58] and deacetylation of PGC-1α, the latter catalysed by Sirt1, with the effects of AMPK on Sirt1 possibly being mediated by increases in NAD^+. AMPK activation also switches off almost all anabolic pathways, including the synthesis of lipids, polysaccharides, ribosomal RNA and proteins (Fig. 4). Finally, turning to regulation of protein synthesis, mTORC1 (which is activated by amino acid availability, as well as by the PI(3)K–Akt pathway) phosphorylates several downstream targets that activate protein synthesis. As already discussed, mTORC1 activation particularly enhances translation of 5′-TOP mRNAs[31] (which include mRNAs encoding ribosomal proteins and translation elongation factors) as well as HIF-1α. AMPK inactivates mTORC1 by two complementary mechanisms[59,60]. Thus, not only does AMPK switch off translation of ribosomal proteins (complementing its effect to switch off ribosomal RNA transcription), but it also downregulates expression of HIF-1α. Supporting the latter, HIF-1α protein expression is markedly upregulated in LKB1-deficient or AMPK-deficient mouse embryo fibroblasts, an effect sensitive to rapamycin[61]. The important role of AMPK in regulating mTORC1 in T cells was also revealed by studies of AMPK-null T cells, which displayed increased glycolysis and expression of glycolytic enzymes, elevated phosphorylation of mTORC1 targets, and release of IFN-γ in response to stimulation of the antigen receptor, with the last two effects being reversed by rapamycin[62].

AMPK as a suppressor of tumorigenesis and inflammation

AMPK activation not only causes cell-cycle arrest but also opposes most of the metabolic changes that occur in rapidly proliferating cells. It therefore seems likely that tumour cells would be under selection pressure in favour of changes that downregulate AMPK. In fact, several such changes have been reported. For example, although loss of — or mutations in — the AMPK genes themselves seem to be relatively rare in tumour cells, genetic loss in *LKB1* is frequent. The *LKB1* gene undergoes homozygous loss-of-function mutations, which are known to abolish activation of AMPK in response to energy stress[43], in around 30% of non-small-cell lung carcinomas[63] and in 20% of cervical cancers.

AMPK is activated by many existing drugs or xenobiotics, including 2-deoxyglucose; AICA riboside (also known as acadesine)[64]; metformin[65], used for treatment of type 2 diabetes; berberine[66], a natural product used in traditional Chinese medicine; and A-769662 (ref. 67), derived from a high-throughput screen for AMPK activators. These have differing mechanisms of action: AICA riboside is metabolized to the ribotide ZMP that mimics the effects of AMP on AMPK[64], A-769662 binds directly to a site on AMPK distinct from those used by adenine nucleotides, whereas the others — being

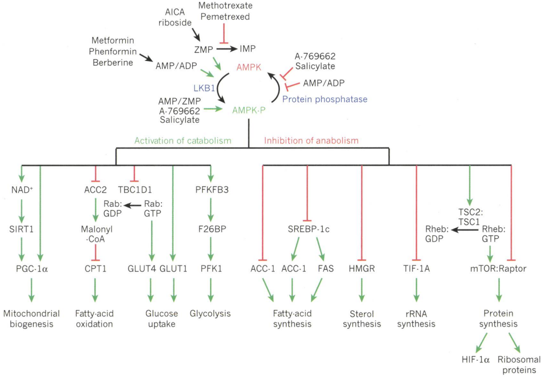

AICA riboside
Methotrexate
Pemetrexed
Metformin
Phenformin
Berberine
ZMP
IMP
AMP/ADP
AMPK
A-769662
Salicylate
AMP/ADP
LKB1
Protein phosphatase
AMP/ZMP
A-769662
Salicylate
AMPK-P
Activation of catabolism
Inhibition of anabolism
NAD+
ACC2
TBC1D1
PFKFB3
Rab: GDP
Rab: GTP
SIRT1
Malonyl-CoA
F26BP
SREBP-1c
TSC2: TSC1
Rheb: GDP
Rheb: GTP
PGC-1α
CPT1
GLUT4
GLUT1
PFK1
ACC-1
ACC-1
FAS
HMGR
TIF-1A
mTOR:Raptor
Mitochondrial biogenesis
Fatty-acid oxidation
Glucose uptake
Glycolysis
Fatty-acid synthesis
Sterol synthesis
rRNA synthesis
Protein synthesis
HIF-1α
Ribosomal proteins

Figure 4 | Regulation of AMPK by drugs, including anti-inflammatory drugs, and the principal metabolic pathways it regulates. AMPK is activated by phosphorylation by LKB1, and inactivated by dephosphorylation by protein phosphatases. Binding of AMP or ADP (which increase in response to drugs inhibiting mitochondrial function; for example, metformin, phenformin and berberine) or ZMP (which increases in response to drugs modulating purine biosynthesis; for example, AICA riboside, methotrexate and pemetrexed) activates AMPK by promoting phosphorylation, by inhibiting dephosphorylation and by direct allosteric activation. A-769662 and the anti-inflammatory agent salicylate bind to AMPK at a site distinct from the nucleotides, causing allosteric activation and inhibition of dephosphorylation. Once activated, AMPK activates catabolic pathways (left) and inhibits anabolic pathways (right). One mechanism to explain how AMPK activation has an anti-inflammatory effects would be its promotion of mitochondrial biogenesis and fatty-acid oxidation (left), while down-regulating glycolysis by inhibiting mTORC1 (shown here as mTOR:Raptor) and hence expression of HIF-1α (right). ACC-1, acetyl-coenzyme A carboxylase-1; HMGR, 3-hydroxy-3-methylglutaryl-coenzyme A reductase; FAS, fatty acid synthase; F26BP, fructose 2,6-bisphosphate; PFK1, phosphofructokinase-1; rRNA, ribosomal RNA; SREBP-1, sterol regulatory element binding protein-1, TIF-1A, transcription intermediary factor-1A; TSC, tuberous sclerosis complex protein.

inhibitors of glycolysis (2-deoxyglucose) or the respiratory chain (metformin and berberine) — activate AMPK indirectly by increasing cellular AMP and ADP[68–71]. Intriguingly, metformin use for the treatment of diabetes is associated with a significantly reduced risk of developing cancer[72].

All of the AMPK activators mentioned have been reported to inhibit inflammatory responses in various model systems. These include inhibition by AICA riboside of iNOS induction by TNF-α, IFN-γ and lipopolysaccharide in myocytes, adipocytes or macrophages grown in culture, or by AICA riboside or metformin in muscle and adipose tissue of lipopolysaccharide-treated rats[73]; prophylactic and therapeutic effects of AICA riboside[74] and metformin[75] in experimental autoimmune encephalomyelitis — a mouse model of multiple sclerosis; inhibition by AICA riboside of lipopolysaccharide-induced NF-κB phosphorylation and TNF-α production in macrophages, differentiation of T_H1 and T_H17 cells *in vitro*, and expression of inflammatory markers in mouse models of acute and chronic colitis[76,77]; inhibition by AICA riboside or metformin production of inflammatory cytokines by primary bronchial epithelial cells from patients with or without cystic fibrosis[78]; inhibition by AICA riboside or berberine of lipopolysaccharide-induced inflammatory responses in neutrophils *in vitro*, and amelioration by AICA riboside of lipopolysaccharide-induced lung injury in mice[79]; reversal of reduced AMPK activation and fatty-acid oxidation in CD8 T cells, and of memory T-cell survival *in vivo*, in mice with defective TRAF6 (ref. 41); and inhibition by 2-deoxyglucose or AICA riboside activation, and the switch to aerobic glycolysis, of mouse dendritic cells induced by lipopolysaccharide; downregulation of AMPK using short hairpin RNA had the opposite effect[6].

It should be emphasized that AMPK activators can have 'off-target' effects, and there is evidence that AICA riboside can have anti-inflammatory effects in primary rat microglial cells that are AMPK-independent[80]. It is therefore important to back up pharmacological studies using genetic approaches. Several such studies do support the anti-inflammatory actions of AMPK: experimental autoimmune encephalomyelitis was more severe in *Ampk-α1*$^{-/-}$ mice than in controls[81]; cultured T cells isolated from spleen of *Ampk-α1*$^{-/-}$ mice secreted more IFN-γ and less IL-17 in response to antigen challenge than those from controls, suggesting that AMPK loss favours a pro-inflammatory state[81]; knockdown of the gene that encodes AMPK in macrophages, or expression of a dominant negative mutant, promoted a pro-inflammatory state[82]; and AMPK-null T cells were pro-inflammatory, hyperactivated and produced more IFN-γ than controls[62].

Relatively few studies have been carried out on the role of AMPK in the metabolism of cells of the immune system. However, many studies have suggested that AMPK downregulates NF-κB signalling through multiple mechanisms, although in most cases this was not in cells of the immune system (reviewed in ref. 83). In mouse bone-marrow dendritic cells, lipopolysaccharide caused a marked reduction of Thr 172 phosphorylation on AMPK, whereas knockdown of AMPK enhanced lipopolysaccharide-induced activation, assessed by expression of interleukin-12 p40 subunit (IL-12p40) and CD86. Administration of AICA riboside reduced expression of IL-12p40 and decreased the switch to aerobic glycolysis, whereas IL-10, an inhibitor of dendritic cell activation, reduced the effect of lipopolysaccharide on AMPK phosphorylation[6]. Glycolysis, and expression of glycolytic enzymes, is also elevated in T cells from *Ampk-α1*$^{-/-}$ mice[62], consistent with the idea that AMPK represses aerobic glycolysis. However, importantly, although the T cells from *Ampk-α1*$^{-/-}$ mice exhibited elevated glycolysis, they did not display a bias towards pro-inflammatory $CD4^+$ lineages. Enhanced glycolysis is therefore necessary, but not sufficient, for such a bias.

Interestingly, there is one situation in which AMPK activation may have a pro-Warburg effect. After induction by lipopolysaccharide, monocytes express the PFKFB3 variant of 6-phosphofructo-2-kinase. AMPK phosphorylates this protein and increases its kinase activity, thus increasing the cellular concentration of the glycolytic activator fructose-2,6-bisphosphate[84]. This mechanism may allow activated macrophages to continue to generate ATP even in hypoxic areas of infection or tissue damage, and should perhaps be viewed as an aspect of the catabolic, ATP-generating function of AMPK.

Activation of AMPK by anti-inflammatory drugs

As already described, several AMPK-activating drugs do have anti-inflammatory actions in animal models, although they are not yet used clinically for this purpose. In addition, the classic anti-inflammatory drug salicylate has recently been found to be an AMPK activator[85]. Salicylate or its derivatives, such as salicin and methyl salicylate, are produced by plants as hormones that are released in response to infection by fungal or bacterial pathogens, and trigger defence responses in uninfected tissues (they can therefore be regarded as equivalent to cytokines). Plant tissues rich in salicylates, such as willow bark, were described as medicines in the third millennium BC, and aspirin (acetylsalicylic acid) is a modern synthetic derivative. Aspirin is rapidly hydrolysed to salicylate following its absorption, with the peak plasma concentration and half-life of salicylate being much higher and longer than those of aspirin. The classic targets for aspirin are the cyclooxygenases involved in prostanoid biosynthesis, which are inhibited potently and irreversibly by aspirin, but less potently (and reversibly) by salicylate[86]. Although aspirin may be responsible for rapid antithrombotic effects owing to the inhibition of thromboxane synthesis in platelets, its more long-term anti-inflammatory effects may instead be mediated by salicylate[86]. It is also not certain that cyclooxygenases are the sole targets for the anti-inflammatory actions of aspirin.

In recent studies, salicylate has been found to activate AMPK by direct binding to the same site as the synthetic activator A-769662, causing conformational changes that both inhibited Thr 172 dephosphorylation and caused allosteric activation[85]. Only AMPK-β1 complexes are sensitive to salicylate and A-769662, and using mice with the gene encoding the β1 complex knocked out it was shown that at least one effect of these drugs *in vivo* — activation of fatty-acid oxidation — was mediated by AMPK[85]. Although the extent to which the anti-inflammatory effects of salicylate are mediated by AMPK remains unclear, the fact that increased fatty-acid oxidation is a common feature of anti-inflammatory cells suggests that this is worth investigating.

Aspirin and salicylate also inhibit the NF-κB pathway, and it has been proposed that this is due to inhibition of the upstream kinase IKK-β (ref. 87). However, inhibition of IKK-β — being competitive with ATP — is much less potent at physiological ATP concentrations[87]. This raises the possibility that inhibition of the NF-κB pathway may be mediated by AMPK activation, rather than by direct inhibition of IKK-β.

Intriguingly, methotrexate — another drug used to treat inflammatory conditions — can also activate AMPK. ZMP, which is synthesized intracellularly from AICA riboside, is an intermediate in purine nucleotide biosynthesis. ZMP is metabolized in two steps to inosine monophosphate (IMP), the first step being catalysed by an enzyme that transfers a formyl group from N10-formyltetrahydrofolate to ZMP. Methotrexate is a folate analogue that inhibits this enzyme, along with other enzymes that utilize tetrahydrofolates. Thus, methotrexate has been found to cause increased accumulation of ZMP and potentiates AMPK activation when breast cancer cells are incubated with AICA riboside[88]. In a human leukaemia cell line, another antifolate drug, pemetrexed, causes AMPK activation even in the absence of AICA riboside[89]. This raises the possibility that the anti-inflammatory effects of antifolate drugs may also be partly mediated by AMPK.

Anti-inflammatory agents may create pseudo-starvation

The ancestral role of AMPK — based on the function of the yeast orthologue — seems to have been in the response to glucose starvation; AMPK is still activated by glucose deprivation in mammalian cells, including cytotoxic T lymphocytes[90]. We have argued in this Review that activation of AMPK in cells of the immune system would promote the switch from a pro-inflammatory to an anti-inflammatory phenotype, partly by causing a switch away from rapid glucose uptake and glycolysis towards mitochondrial oxidative metabolism, including fatty-acid oxidation. Intriguingly, the AMPK-activating drug metformin imposes this switch on cells that carry the CD4 antigen[30]. In cells that are growing or proliferating rapidly, including activated T lymphocytes, aerobic glycolysis allows the use of the TCA cycle to provide precursors for biosynthesis, rather than purely for ATP production. A reduced reliance on the TCA cycle and the respiratory chain for ATP production also allows mitochondria to generate ROS, which are used in antibacterial defence, and to accumulate TCA-cycle intermediates such as citrate that may have a signalling role in inflammation. However, the downside of increased use of glycolysis and the pentose phosphate pathway by inflammatory cells is that they will have a greatly increased demand for glucose. Accelerated glucose use by immune cells is a luxury that may not be possible when the individual is in a state of starvation, helping to explain why starvation depresses the immune system, although glucocorticoids undoubtedly also have a role here.

In an over-nourished state, such as during obesity, there may be a shift towards pro-inflammatory states of immune cells, whereas during dietary restriction the balance may shift towards anti-inflammatory phenotypes instead. Interestingly, in the nematode *Caenorhabditis elegans*, the orthologues of Sirt1 and AMPK, both of which seem to promote anti-inflammatory states in mammals, are required for extension of lifespan in response to dietary restriction[91,92]. The plant polyphenol resveratrol, which extends lifespan in *C. elegans* and in mice that are fed a high-fat diet, has also been reported to activate both Sirt1 and AMPK, although which of the two is the primary target remains controversial[93]. As already discussed, at least two other existing drugs used as anti-inflammatory agents (salicylate and methotrexate) also activate AMPK. Finally, rapamycin can also create a pseudo-starvation state, promoting fatty-acid oxidation and generating T_{reg} cells by inhibiting mTORC1, and inhibiting dendritic cell activation by lipopolysaccharide. It is therefore possible that many anti-inflammatory agents work, in part, by hijacking a system that decreases inflammation in the face of nutrient deprivation. Given the similarities in the metabolism of tumour cells and pro-inflammatory cells, it is not surprising that the same agents — many of which are AMPK activators — should have both antitumour and anti-inflammatory effects. In this context, there is epidemiological evidence that patients who take aspirin long term seem to have some protection against cancer[94]. It is also conceivable that some of the protection against cancer in people with type 2 diabetes provided by the AMPK-activating drug metformin[95] could be due to systemic effects on inflammation.

In addition to glucocorticoids, fasting and starvation in humans cause changes in the circulating levels of many other hormones or adipokines. Interestingly, many of the hormones associated with an overfed state (such as insulin and leptin) inhibit AMPK[96], whereas those associated with an under-fed state (such as ghrelin and adiponectin) activate AMPK[97]. Thus, it seems appropriate that this ancient starvation signalling pathway should have a role in the depression of inflammation, a bodily response that is very expensive in terms of glucose consumption.

Final perspectives

Metabolism has recently returned with a vengeance to become a hot topic in biomedical research. Insight into the changes in metabolism that occur in metabolic diseases, cancer and inflammation may allow us to understand more about disease pathogenesis and point the way to new therapies. Multiple pathways and processes have been uncovered that might lend themselves to therapeutic manipulation. The mechanistic overlap in the metabolic changes that are evident in cancer and inflammation could indicate that some of the targets they have in common (for example, AMPK, mTORC1 or HIF-1α) could be useful in treatment of both disorders. Finally, a reconfiguring of our understanding of metabolism is needed to provide new approaches. Both tumour cells and inflammatory cells are often viewed as acting in a cell-autonomous manner that ignores the needs of the whole organism. However, a key feature that emerged from studies during the heyday of metabolism research was that metabolism is a closely integrated system, involving many hormones that coordinate functions between different tissues. Perhaps the insulin resistance that occurs during inflammatory conditions reflects one aspect of this, as it would reduce glucose uptake by tissues such as muscle, liver and adipocytes, conserving glucose for use by macrophages and T_H17 cells instead. On the other hand, chronic, uncontrolled inflammation or cancer may deplete other cells and tissues of nutrients, promoting cachexia. Perhaps this is why chronic inflammation is so damaging, and shortens lifespan in the same way as cancer. It is also revealing that inflamed tissues, such as the rheumatoid synovium, can be visualized with $[^{18}F]$-fluorodeoxyglucose[98], reflecting their increased glucose uptake in the same way as tumours. This might help to explain why inflammation lies at the heart of so many diseases — it puts the whole body under metabolic stress, driving symptoms and causing morbidity. The prospect of targeting the altered metabolism in inflammation, as well as in cancer, holds substantial therapeutic promise. ■

Received 1 August; accepted 13 November 2012.

1. Warburg, O. Metabolism of tumours. *Biochem. Z.* **142,** 317–333 (1923).
2. Oren, P., Farnham, A. E., Milofsky, M. & Marnovsky, M. L. Metabolic patterns in three types of phagocytizing cells. *J. Cell Biol.* **17,** 487–501 (1963).
3. Bennett, W. E. & Cohn, Z. A. The isolation and selected properties of blood monocytes. *J. Exp. Med.* **123,** 145–159 (1966).
4. Arwadi, M. S. & Newsholme, E. A. Maximum activities of some enzymes of glycolysis, the CS cycle and ketone body and glutamine utilization pathways in lymphocytes of the rat. *Biochem. J.* **208,** 743–748 (1982).
5. Newsholme, P., Curi, R., Gordon, S. & Newsholme, E. A. Metabolism of glucose, glutamine, long-chain fatty acids and ketone bodies by murine macrophages. *Biochem. J.* **239,** 121–125 (1986).
6. Krawczyk, C. *et al.* Toll-like receptor-induced changes in glycolytic metabolism regulate dendritic cell activation. *Blood* **115,** 4742–4749 (2010).
   This paper describes for the first time how lipopolysaccharide induces aerobic glycolysis in dendritic cells.
7. Rodriguez-Prados, J. C. *et al.* Substrate fate in activated macrophages: a comparison between innate, classic and alternative activation. *J. Immunol.* **185,** 605–614 (2010).
8. Shi, L. Z. *et al.* HIF-1α-dependent glycolytic pathway orchestrates a metabolic checkpoint for the differentiation of T_H17 and T_{reg} cells. *J. Exp. Med.* **208,** 1367–1379 (2011).
9. Dang, E. V. *et al.* Control of T_H17/T_{reg} balance by hypoxia-inducible factor 1. *Cell* **146,** 772–784 (2012).
   References 8 and 9 provide a detailed insight into the role of HIF-1α in T_H17 cells for the promotion of glycolysis and induction of RORγt.
10. van der Windt, G. J. W. *et al.* Mitochondrial respiratory capacity is a critical regulator of $CD8^+$ T cell memory development. *Immunity* **36,** 68–78 (2012).
    This paper describes the predominance of mitochondrial metabolism in memory T lymphocytes.
11. Wang, R. *et al.* The transcription factor Myc controls metabolic reprogramming upon T lymphocyte activation. *Immunity* **35,** 871–882 (2011).
12. Everts, B. *et al.* Commitment to glycolysis sustains survival of nitric oxide-producing inflammatory dendritic cells. *Blood* **120,** 1422–1431 (2012).
13. Garedew, A., Henderson, S. O. & Moncada, S. Activated macrophages utilize glycolytic ATP to maintain mitochondrial membrane potential and prevent apoptotic death. *Cell Death Differ.* **17,** 1540–1550 (2010).
14. West, A. P. *et al.* TLR signalling augments macrophage bactericidal activity through mitochondrial ROS. *Nature* **472,** 476–482 (2011).
    A key role for mitochondrial ROS activated by TLR4 in bactericidal activity is described.
15. Sonoda, J. *et al.* Nuclear receptor ERRα and coactivator PGC-1β are effectors of IFN-γ-induced host defense. *Genes Dev.* **21,** 1909–1920 (2007).
16. Vats, D. *et al.* Oxidative metabolism and PGC-1β attenuates macrophage-mediated inflammation. *Cell Metab.* **4,** 13–24 (2006).
17. Wu, Z. *et al.* Mechanisms controlling mitochondrial biogenesis and respiration through the thermogenic co-activator PGC-1. *Cell* **98,** 115–124 (2010).
18. Schwer, B. & Verdin, E. Conserved metabolic regulatory functions of sirtuins. *Cell Metab.* **7,** 104–112 (2008).
19. Liu, T. F., Vachharajani, V. T., Yoza, B. K. & McCall, C. E. NAD^+-dependent sirtuin 1 and 6 proteins coordinate a switch from glucose to fatty acid oxidation during the acute inflammatory response. *J. Biol. Chem.* **287,** 25758–25769 (2012).
20. Yeung, F. *et al.* Modulation of NF-κB-dependent transcription and cell survival by Sirt1 deacetylase. *EMBO J.* **23,** 2369–2380 (2004).
21. Kelly, T. J., Lerin, C., Haas, W., Gygi, S. P. & Puigserver, P. GCN5-mediated transcriptional control of the metabolic coactivator PGC-1β through lysine acetylation. *J. Biol. Chem.* **284,** 19945–19952 (2009).
22. Lin, J., Handschin, C. & Spiegelman, B. M. Metabolic control through the PGC-1 family of transcription coactivators. *Cell Metab.* **1,** 361–370 (2005).

23. Liu, T. F. *et al.* NAD$^+$-dependent SIRT1 deacetylase participates in epigenetic reprogramming during endotoxin tolerance. *J. Biol. Chem.* **286,** 9856–9864 (2011).
24. Haschemi, A. *et al.* The sedoheptulose kinase CARKL directs macrophage polarization through control of glucose metabolism. *Cell Metab.* **15,** 813–826 (2012).
    This paper describes the role of CARKL in M2 macrophage function.
25. Nishi, T. *et al.* Spatial redox regulation of a critical cysteine residue in NF-κB *in vivo*. *J. Biol. Chem.* **277,** 44548–44556 (2002).
26. Newsholme, P. & Newsholme, E. A. Rates of utilization of glucose, glutamine and oleate and formation of end-products by mouse peritoneal macrophages in culture. *Biochem. J.* **261,** 211–218 (1989).
27. Infantino, V. *et al.* The mitochondrial citrate carrier: a new player in inflammation. *Biochem. J.* **438,** 433–436 (2011).
28. Hedesov, C. J. & Esmann, V. Respiration and glycolysis of normal human lymphocytes. *Blood* **28,** 163–174 (1966).
29. Frauwirth K. A. *et al.* The CD28 signalling pathway regulates glucose metabolism. *Immunity* **16,** 769–777 (2002).
30. Michalek, R. D. *et al.* Distinct glycolytic and lipid oxidative metabolic programs are essential for effector and regulatory CD4$^+$ T cell subsets. *J. Immunol.* **186,** 3299–3303 (2011).
    This paper gives the first detailed description of the metabolic features of effector and regulatory T cells.
31. Huo, Y. *et al.* Stable isotope-labelling analysis of the impact of inhibition of the mammalian target of rapamycin on protein synthesis. *Biochem. J.* **444,** 141–151 (2012).
32. Duvel, K. *et al.* Activation of a metabolic gene regulatory network downstream of mTOR complex. *Mol. Cell* **39,** 171–183 (2010).
33. Powell, J. D. *et al.* Regulation of immune responses by mTOR. *Annu. Rev. Immunol.* **30,** 39–68 (2011).
34. Delgoffe, G. M. *et al.* The mTOR kinase differentially regulates effector and regulatory T cell lineage commitment. *Immunity* **30,** 832–844 (2009).
35. Battaglia, M. *et al.* Rapamycin selectively expands CD4$^+$CD25$^+$Foxp3$^+$ regulatory T cells. *Blood* **105,** 4743–4748 (2005).
36. Degoffe, G. M. *et al.* The kinase mTOR regulates the differentiation of helper T cells through the selective activation of signalling by mTORC1 and mTORC2. *Nature Immunol.* **12,** 295–303 (2011).
    This article highlights the role of mTOR in metabolic regulation in T$_H$ cells.
37. Peyssonnaux, C. *et al.* Essential role of hypoxia inducible factor-1α in development of lipopolysaccharide-induced sepsis. *J. Immunol.* **178,** 7516–7519 (2007).
38. Nizet, V. & Johnson, R. S. Interdependence of hypoxic and innate immune responses. *Nature Rev. Immunol.* **9,** 609–617 (2009).
39. Cramer, T. *et al.* HIF-1α is essential for myeloid cell-mediated inflammation. *Cell* **112,** 645–657 (2003).
    This paper describes the crucial function of HIF-1α in inflammation.
40. Imtiyaz, H. Z. *et al.* Hypoxia-inducible factor 2α regulates macrophage function in mouse models of acute and tumor inflammation. *J. Clin. Invest.* **120,** 2699–2714 (2010).
41. Pearce, E. L. *et al.* Enhancing CD8 T cell memory by modulating fatty acid oxidation. *Nature* **460,** 103–107 (2009).
    This article is the first report on the key role of fatty-acid oxidation in mitochondria in memory T cells.
42. Araki, K. *et al.* mTOR regulates memory CD8 T-cell differentiation. *Nature* **460,** 108–112 (2009).
43. Hawley, S. A. *et al.* Complexes between the LKB1 tumor suppressor, STRADα/β and MO25α/β are upstream kinases in the AMP-activated protein kinase cascade. *J. Biol.* **2,** 28 (2003).
44. Woods, A. *et al.* LKB1 is the upstream kinase in the AMP-activated protein kinase cascade. *Curr. Biol.* **13,** 2004–2008 (2003).
45. Shaw, R. J. *et al.* The tumor suppressor LKB1 kinase directly activates AMP-activated kinase and regulates apoptosis in response to energy stress. *Proc. Natl Acad. Sci. USA* **101,** 3329–3335 (2004).
46. Alessi, D. R., Sakamoto, K. & Bayascas, J. R. Lkb1-dependent signaling pathways. *Annu. Rev. Biochem.* **75,** 137–163 (2006).
47. Lizcano, J. M. *et al.* LKB1 is a master kinase that activates 13 protein kinases of the AMPK subfamily, including the MARK/PAR-1 kinases. *EMBO J.* **23,** 833–843 (2004).
48. Hawley, S. A. *et al.* 5′-AMP activates the AMP-activated protein kinase cascade, and Ca^{2+}/calmodulin the calmodulin-dependent protein kinase I cascade, via three independent mechanisms. *J. Biol. Chem.* **270,** 27186–27191 (1995).
49. Davies, S. P., Helps, N. R., Cohen, P. T. W. & Hardie, D. G. 5′-AMP inhibits dephosphorylation, as well as promoting phosphorylation, of the AMP-activated protein kinase. Studies using bacterially expressed human protein phosphatase-2Cα and native bovine protein phosphatase-2AC. *FEBS Lett.* **377,** 421–425 (1995).
50. Oakhill, J. S. *et al.* AMPK is a direct adenylate charge-regulated protein kinase. *Science* **332,** 1433–1435 (2011).
51. Woods, A. *et al.* Yeast SNF1 is functionally related to mammalian AMP-activated protein kinase and regulates acetyl-CoA carboxylase *in vivo*. *J. Biol. Chem.* **269,** 19509–19515 (1994).
52. Wilson, W. A., Hawley, S. A. & Hardie, D. G. The mechanism of glucose repression/derepression in yeast: SNF1 protein kinase is activated by phosphorylation under derepressing conditions, and this correlates with a high AMP:ATP ratio. *Curr. Biol.* **6,** 1426–1434 (1996).
53. Gancedo, J. M. Carbon catabolite repression in yeast. *Eur. J. Biochem.* **206,** 297–313 (1992).
54. Haurie, V., Sagliocco, F. & Boucherie, H. Dissecting regulatory networks by means of two-dimensional gel electrophoresis: application to the study of the diauxic shift in the yeast *Saccharomyces cerevisiae*. *Proteomics* **4,** 364–373 (2004).
55. Imamura, K., Ogura, T., Kishimoto, A., Kaminishi, M. & Esumi, H. Cell cycle regulation via p53 phosphorylation by a 5′-AMP activated protein kinase activator, 5-aminoimidazole-4-carboxamide-1-β-D-ribofuranoside, in a human hepatocellular carcinoma cell line. *Biochem. Biophys. Res. Commun.* **287,** 562–567 (2001).
56. Jones, R. G. *et al.* AMP-activated protein kinase induces a p53-dependent metabolic checkpoint. *Mol. Cell* **18,** 283–293 (2005).
57. Lin, J., Handschin, C. & Spiegelman, B. M. Metabolic control through the PGC-1 family of transcription coactivators. *Cell Metab.* **1,** 361–370 (2005).
58. Jager, S., Handschin, C., St-Pierre, J. & Spiegelman, B. M. AMP-activated protein kinase (AMPK) action in skeletal muscle via direct phosphorylation of PGC-1α. *Proc. Natl Acad. Sci. USA* **104,** 12017–12022 (2007).
59. Inoki, K., Zhu, T. & Guan, K. L. TSC2 mediates cellular energy response to control cell growth and survival. *Cell* **115,** 577–590 (2003).
60. Gwinn, D. M. *et al.* AMPK phosphorylation of raptor mediates a metabolic checkpoint. *Mol. Cell* **30,** 214–226 (2008).
61. Shackelford, D. B. *et al.* mTOR and HIF-1α-mediated tumor metabolism in an LKB1 mouse model of Peutz-Jeghers syndrome. *Proc. Natl. Acad. Sci. USA* **106,** 11137–11142 (2009).
62. MacIver, N. J. *et al.* The liver kinase B1 is a central regulator of T cell development, activation, and metabolism. *J. Immunol.* **187,** 4187–4198 (2011).
63. Sanchez-Cespedes, M. *et al.* Inactivation of LKB1/STK11 is a common event in adenocarcinomas of the lung. *Cancer Res.* **62,** 3659–3662 (2002).
64. Corton, J. M., Gillespie, J. G., Hawley, S. A. & Hardie, D. G. 5-Aminoimidazole-4-carboxamide ribonucleoside: a specific method for activating AMP-activated protein kinase in intact cells? *Eur. J. Biochem.* **229,** 558–565 (1995).
65. Zhou, G. *et al.* Role of AMP-activated protein kinase in mechanism of metformin action. *J. Clin. Invest.* **108,** 1167–1174 (2001).
66. Lee, Y. S. *et al.* Berberine, a natural plant product, activates AMP-activated protein kinase with beneficial metabolic effects in diabetic and insulin-resistant states. *Diabetes* **55,** 2256–2264 (2006).
67. Cool, B. *et al.* Identification and characterization of a small molecule AMPK activator that treats key components of type 2 diabetes and the metabolic syndrome. *Cell Metab.* **3,** 403–416 (2006).
68. Owen, M. R., Doran, E. & Halestrap, A. P. Evidence that metformin exerts its anti-diabetic effects through inhibition of complex 1 of the mitochondrial respiratory chain. *Biochem. J.* **348,** 607–614 (2000).
69. El-Mir, M. Y. *et al.* Dimethylbiguanide inhibits cell respiration via an indirect effect targeted on the respiratory chain complex I. *J. Biol. Chem.* **275,** 223–228 (2000).
70. Goransson, O. *et al.* Mechanism of action of A-769662, a valuable tool for activation of AMP-activated protein kinase. *J. Biol. Chem.* **282,** 32549–32560 (2007).
71. Hawley, S. A. *et al.* Use of cells expressing γ-subunit variants to identify diverse mechanisms of AMPK activation. *Cell Metab.* **11,** 554–565 (2010).
72. Soranna, D. *et al.* Cancer risk associated with use of metformin and sulfonylurea in type 2 diabetes: a meta-analysis. *Oncologist* **17,** 813–822 (2012).
73. Pilon, G., Dallaire, P. & Marette, A. Inhibition of inducible nitric-oxide synthase by activators of AMP-activated protein kinase: a new mechanism of action of insulin-sensitizing drugs. *J. Biol. Chem.* **279,** 20767–20774 (2004).
74. Nath, N. *et al.* 5-Aminoimidazole-4-carboxamide ribonucleoside: a novel immunomodulator with therapeutic efficacy in experimental autoimmune encephalomyelitis. *J. Immunol.* **175,** 566–574 (2005).
    This is the first report to give a detailed account of the anti-inflammatory effects of the AMPK activator AICA riboside.
75. Nath, N. *et al.* Metformin attenuated the autoimmune disease of the central nervous system in animal models of multiple sclerosis. *J. Immunol.* **182,** 8005–8014 (2009).
76. Bai, A. *et al.* AMPK agonist downregulates innate and adaptive immune responses in TNBS-induced murine acute and relapsing colitis. *Biochem. Pharmacol.* **80,** 1708–1717 (2010).
77. Bai, A. *et al.* Novel anti-inflammatory action of 5-aminoimidazole-4-carboxamide ribonucleoside with protective effect in dextran sulfate sodium-induced acute and chronic colitis. *J. Pharmacol. Exp. Ther.* **333,** 717–725 (2010).
78. Myerburg, M. M. *et al.* AMPK agonists ameliorate sodium and fluid transport and inflammation in cystic fibrosis airway epithelial cells. *Am. J. Respir. Cell Mol. Biol.* **42,** 676–684 (2010).
79. Zhao, X. *et al.* Activation of AMPK attenuates neutrophil proinflammatory activity and decreases the severity of acute lung injury. *Am. J. Physiol. Lung Cell. Mol. Physiol.* **295,** L497–L504 (2008).
80. Labuzek, K., Liber, S., Gabryel, B. & Okopien, B. AICAR (5-aminoimidazole-4-carboxamide-1-β-4-ribofuranoside) increases the production of toxic molecules and affects the profile of cytokines release in LPS-stimulated rat primary microglial cultures. *Neurotoxicology* **31,** 134–146 (2010).
81. Nath, N. *et al.* Loss of AMPK exacerbates experimental autoimmune encephalomyelitis disease severity. *Biochem. Biophys. Res. Comm.* **386,** 16–20 (2009).
82. Sag, D., Carling, D., Stout, R. D. & Suttles, J. Adenosine 5′-monophosphate-activated protein kinase promotes macrophage polarization to an anti-inflammatory functional phenotype. *J. Immunol.* **181,** 8633–8641 (2008).
83. Salminen, A., Hyttinen, J. M. & Kaarniranta, K. AMP-activated protein kinase inhibits NF-κB signaling and inflammation: impact on healthspan and lifespan. *J. Mol. Med.* **89,** 667–676 (2011).

84. Marsin, A. S., Bouzin, C., Bertrand, L. & Hue, L. The stimulation of glycolysis by hypoxia in activated monocytes is mediated by AMP-activated protein kinase and inducible phosphofructo-2-kinase. *J. Biol. Chem.* **277,** 30778–30783 (2002).
85. Hawley, S. A. *et al.* The ancient drug salicylate directly activates AMP-activated protein kinase. *Science* **336,** 918–922 (2012).
This paper describes how salicylates might mediate their anti-inflammatory effects through AMPK activation.
86. Higgs, G. A., Salmon, J. A., Henderson, B. & Vane, J. R. Pharmacokinetics of aspirin and salicylate in relation to inhibition of arachidonate cyclooxygenase and anti-inflammatory activity. *Proc. Natl Acad. Sci. USA* **84,** 1417–1420 (1987).
87. Yin, M. J., Yamamoto, Y. & Gaynor, R. B. The anti-inflammatory agents aspirin and salicylate inhibit the activity of IκB kinase-β. *Nature* **396,** 77–80 (1998).
88. Beckers, A. *et al.* Methotrexate enhances the anti-anabolic and antiproliferative effects of 5-aminoimidazole-4-carboxamide riboside. *Mol. Cancer Ther.* **5,** 2211–2217 (2006).
89. Racanelli, A. C., Rothbart, S. B., Heyer, C. L. & Moran, R. G. Therapeutics by cytotoxic metabolite accumulation: pemetrexed causes ZMP accumulation, AMPK activation, and mammalian target of rapamycin inhibition. *Cancer Res.* **69,** 5467–5474 (2009).
90. Rolf, J., Zarrouk, M., Finlay, D. K., Foretz, M., Viollet, B., Cantrell, D. A. AMPKα1: a glucose sensor that controls CD8 T-cell memory. *Eur. J. Immunol.* (in the press).
91. Greer, E. L. *et al.* An AMPK–FOXO pathway mediates longevity induced by a novel method of dietary restriction in *C. elegans*. *Curr. Biol.* **17,** 1646–1656 (2007).
92. Baur, J. A. *et al.* Resveratrol improves health and survival of mice on a high-calorie diet. *Nature* **444,** 337–342 (2006).
93. Canto, C. *et al.* AMPK regulates energy expenditure by modulating NAD$^+$ metabolism and SIRT1 activity. *Nature* **458,** 1056–1060 (2009).
94. Rothwell, P. M. *et al.* Short-term effects of daily aspirin on cancer incidence, mortality, and non-vascular death: analysis of the time course of risks and benefits in 51 randomised controlled trials. *Lancet* **379,** 1602–1612 (2012).
95. Soranna, D. *et al.* Cancer risk associated with use of metformin and sulfonylurea in type 2 diabetes: a meta-analysis. *Oncologist* **17,** 813–822 (2012).
96. Horman, S. *et al.* Insulin antagonizes ischemia-induced Thr172 phosphorylation of AMP-activated protein kinase α-subunits in heart via hierarchical phosphorylation of Ser485/491. *J. Biol. Chem.* **281,** 5335–5340 (2006).
97. Yamauchi, T. *et al.* Adiponectin stimulates glucose utilization and fatty-acid oxidation by activating AMP-activated protein kinase. *Nature Med.* **6,** 1288–1295 (2002).
98. Matsui, T. *et al.* Inflammatory cytokines and hypoxia contribute to ^{18}F-FDG uptake by cells involved in pannus formation in rheumatoid arthritis. *J. Nucl. Med.* **50,** 920–926 (2009).

Acknowledgements Work in the D.G.H. laboratory is supported by the Wellcome Trust and Cancer Research UK. Work in the L.O.N. laboratory is supported by Science Foundation Ireland and the European Research Council.

Author Information Reprints and permissions information is available at www.nature.com/reprints. The authors declare no competing financial interests. Readers are welcome to comment on the online version of this article at go.nature.com/glj9u2. Correspondence should be addressed to L.A.J.O. (laoneill@tcd.ie).

doi:10.1038/nature11863

Fanconi anaemia and the repair of Watson and Crick DNA crosslinks

Molly C. Kottemann[1] & Agata Smogorzewska[1]

The function of Fanconi anaemia proteins is to maintain genomic stability. Their main role is in the repair of DNA interstrand crosslinks, which, by covalently binding the Watson and the Crick strands of DNA, impede replication and transcription. Inappropriate repair of interstrand crosslinks causes genomic instability, leading to cancer; conversely, the toxicity of crosslinking agents makes them a powerful chemotherapeutic. Fanconi anaemia proteins can promote stem-cell function, prevent tumorigenesis, stabilize replication forks and inhibit inaccurate repair. Recent advances have identified endogenous aldehydes as possible culprits of DNA damage that may induce the phenotypes seen in patients with Fanconi anaemia.

Fanconi anaemia occurs in about 1 in every 100,000 births[1]. Biallelic mutations in Fanconi anaemia genes lead to bone-marrow failure and susceptibility to both acute myeloid leukaemia (AML) and solid tumours, congenital abnormalities and infertility[2]. The function of the pathway, on the other hand, is anything but rarefied: Fanconi anaemia proteins participate in the repair of extraordinarily deleterious lesions, interstrand crosslinks, and in maintaining genomic stability during DNA replication.

Fanconi anaemia is genetically heterogeneous. So far, 15 genes have been identified as mutated in patients (Fig. 1 and Table 1), and many more interacting genes have been discovered. Furthermore, there are still patients in whom a mutation has yet to be identified. The known Fanconi anaemia genes work together in interstrand crosslink repair to coordinate the actions of multiple repair processes, in particular nucleases that are necessary for cutting out the interstrand crosslink and for other nucleolytic processing needed for repair, trans-lesion synthesis (TLS) (a mode of damage tolerance that uses specialized polymerases to insert a base across from a lesion or abasic site) and homologous recombination (the pathway best known for its role in repairing double-strand breaks). Fanconi anaemia proteins may also counteract some of the activities of the non-homologous end joining (NHEJ) pathway, an error-prone repair pathway that is used to directly religate DNA ends. Overall, the Fanconi anaemia pathway interacts with many of the genome maintenance pathways, and studying it provides a unique window into the elaborate interplay of multiple cellular networks.

As new functions come to light, the list of biological processes in which the Fanconi anaemia proteins are intimately involved is expanding (Fig. 2). The bone-marrow failure seen in patients with the disorder suggests a function for the pathway in stem-cell biology, and recent work has begun to gain tantalizing insight into not only Fanconi anaemia pathogenesis, but also the intersection between DNA repair and stem-cell maintenance. Several genes involved in the Fanconi anaemia pathway also predispose individuals to breast and ovarian cancer, which, together with the cancer susceptibility of patients with Fanconi anaemia, suggests Fanconi anaemia proteins have an important role in suppressing tumorigenesis. Recent work has also described an unanticipated role for Fanconi anaemia proteins in replication-fork stabilization. Finally, the identification of aldehydes as endogenous genotoxins in Fanconi anaemia provides an exciting glimpse into the challenges in the cellular environment that are important in human disease. In this Review, we highlight these topics and discuss the current understanding of the molecular details of the Fanconi anaemia pathway.

Stem-cell function in Fanconi anaemia pathway absence

One of the clinical identifiers of Fanconi anaemia is bone-marrow failure, which can be treated by bone-marrow transplantation, suggesting that disrupting the Fanconi anaemia pathway leads to dysfunction of the haematopoietic stem cells and progenitor cells themselves. Although the bulk of Fanconi anaemia research so far has concentrated on the DNA repair functions of the pathway in terminally differentiated cells, this clinical manifestation points to a crucial role in stem-cell development that must be explored. A clear connection between the cellular and disease phenotypes in Fanconi anaemia has been elusive, owing to the difficulty in establishing mouse models that parallel human disease and because the limited proliferative potential of primary haematopoietic stem cells makes patient-derived *in vitro* systems unfeasible. Understanding the interactions between the function of the Fanconi anaemia pathway and stem-cell regulation and maintenance will probably prove to be key to elucidating the functional link between the pathway and the disease.

Recent research provides a glimpse into a functional link between defects in genome maintenance and bone-marrow failure in Fanconi anaemia. Levels of p53, one of the central 'guardians' of the genome, and its transcriptional target p21 are elevated in primary blood, bone marrow and even in the livers of fetuses with Fanconi anaemia during haematopoietic stem-cell expansion in the organ. Perhaps because of this hyperactivation, the pool of available haematopoietic stem cells and progenitor cells is already compromised at birth in individuals with Fanconi anaemia, and the bone marrow of these patients is less proliferative than the bone marrow of healthy individuals — with more cells in the G0 and G1 cell-cycle stages and an increase in DNA damage signalling, as indicated by the presence of phosphorylated histone H2AX. Not only are there fewer haematopoietic stem cells and progenitor cells in Fanconi anaemia bone marrow, but those stem cells are also unable to produce progenitors *in vitro*[3].

Disrupting the Fanconi anaemia pathway is itself sufficient to impair haematopoietic development *in vitro*. Short-hairpin RNA knockdown of either *FANCA* or *FANCD2* in human embryonic stem cells leads to significantly reduced production of haematopoietic stem and progenitor cells and their differentiated daughter populations[4]. Exposing progenitor cells to crosslinking agents exacerbates the proliferation deficiency. These findings link the canonical interstrand crosslink repair function of the Fanconi anaemia pathway to its role in stem-cell development. In murine haematopoietic stem cells and human bone-marrow stromal cells, Fanconi anaemia pathway deficiency leads to an increase in binucleated cells, suggesting that cytokinesis failure may also contribute to bone-marrow failure[5].

---

[1] Laboratory of Genome Maintenance, The Rockefeller University, New York 10065, USA.

An important role for the Fanconi anaemia pathway in stem-cell function was uncovered when researchers attempted to derive induced pluripotent stem (iPS) cells from Fanconi anaemia fibroblasts. They discovered that fibroblasts that are deficient for the Fanconi anaemia pathway are refractory to reprogramming[6,7]. It has been proposed, although not shown, that Fanconi anaemia proteins participate directly in reprogramming. Alternatively, the increase in DNA damage signalling in Fanconi anaemia cells may preclude the normal cell-cycle progression (division) necessary for reprogramming. It is also possible that the Fanconi anaemia pathway is required to repair DNA damage that reprogramming itself induces. Discriminating between these possibilities will be key to understanding the role of the Fanconi anaemia pathway in stem cells.

Taken together, recent data suggest a connection between DNA repair defects and haematopoietic stem cell and progenitor cell failure, although many mysteries remain. It is intriguing, for example, that Fanconi anaemia pathway deficiency leads to problems with both quiescent haematopoietic stem cells and their rapidly dividing progenitors — two very different cellular landscapes. An appreciation of the underlying molecular functions of Fanconi anaemia proteins in haematopoietic stem-cell populations will provide a better understanding of the pathophysiology of the disease. Furthermore, teasing out how a pathway that is traditionally associated with DNA repair is important to the normal function of a stem-cell population will provide insight into stem-cell biology in general.

Fanconi anaemia pathway and tumour suppression

Fanconi anaemia has a complicated relationship with tumorigenesis. Although susceptibility to cancer, particularly AML and squamous cell carcinomas, is endemic in patients with homozygous mutations in Fanconi anaemia genes, four out of five downstream members of the pathway — which function in homology-directed repair of DNA ends — also confer susceptibility to breast cancer and ovarian cancer when mutated in only one copy[8–14] (Table 1). The best known breast cancer predisposition gene *BRCA2* (also known as *FANCD1*)[15] is involved in the Fanconi anaemia pathway, as is its partner *PALB2* (also known as *FANCN*)[16], as well as *BRIP1* (also known as *FANCJ*)[17,18]. Defects in *RAD51C*, one of the RAD51 paralogues important for homologous recombination, are the cause of a Fanconi-anaemia-like disorder, and *RAD51C* is provisionally called *FANCO*[19]. Patients with biallelic *BRCA2* and *PALB2* mutations evince more severe phenotypes with more prevalent and earlier onset of leukaemia and appearence of embryonal tumours in the first 2 years of life.

The cancers that develop in patients with Fanconi anaemia are indicative of the complex landscape that is required for tumorigenesis to progress. Predisposition to AML may derive from the same haematopoietic stem-cell instability that contributes to bone-marrow failure. The types of squamous cell carcinoma of the head and neck that are common to patients with Fanconi anaemia are also associated with human papilloma virus, raising the possibility that these tumours develop in the context of infection with an oncogenic virus.

The correlation between downstream members of the Fanconi anaemia pathway and cancer is intuitive when one considers the mechanics of the pathway. The upstream members that belong to the core and FANCI–FANCD2 (I–D2) complex are more self-contained, functioning in pathway activation and early coordination steps, whereas BRCA2 and PALB2 are intimately associated with homology-directed repair (homologous recombination), which is used more broadly outside and independently of the Fanconi anaemia pathway.

Fanconi anaemia proteins in interstrand crosslink repair

Crosslinked DNA impedes both transcription and replication; interstrand crosslinks, therefore, need to be removed during all stages of the cell cycle. Indeed, there is evidence for both replication-dependent and replication-independent repair. In the absence of damage sensing by active replication forks and a homologous template, repair depends on nucleotide excision repair, especially transcription-coupled nucleotide excision repair and TLS proteins (including polymerase κ)[20,21]. The involvement of Fanconi anaemia and Fanconi-anaemia-associated proteins in G1 repair remains to be elucidated, although the core complex components and the I–D2 complex are able to bind to interstrand crosslinks independently of replication and their absence results in inefficient repair[22].

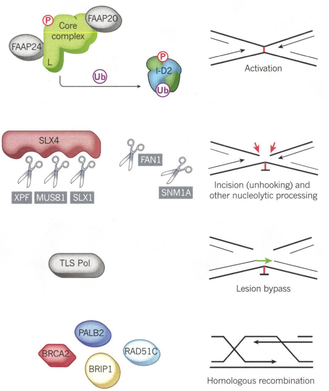


Figure 1 | The Fanconi anaemia pathway and interstrand crosslink repair. The core complex is made up of Fanconi anaemia proteins and accessory proteins (such as FAAP20 and FAAP24, shown here). On detection of the crosslink, the complex is activated by ATR-mediated phosphorylation (P), and one component, FANCL, ubiquitinates (Ub) the I–D2 complex. The I–D2 complex then coordinates the action of downstream repair factors. SLX4 functions as a scaffold for the three nucleases XPF, MUS81 and SLX1 that function at the site of DNA damage to make incisions either side of two covalently linked nucleotides. FAN1 and SNM1A also have a role in processing the crosslink after incision, although the precise mechanism is unclear. On the incised strand, TLS polymerases (TLS Pol) are recruited to bypass the unhooked crosslink. The break is then repaired through homologous recombination involving the Fanconi anaemia proteins BRCA2, BRIP1, PALB2 and RAD51C. SLX4 may also have a role here by regulating the nucleases MUS81 and SLX1. BRIP1 is implicated in homologous recombination but is also required at an earlier step for pathway activation. The proteins shown in colour are Fanconi anaemia proteins. Those in grey are important for interstrand crosslink repair, but their genes have not been found to be mutated in patients with Fanconi anaemia.

During S phase, the replication fork encounters a lesion and is forced to stall, activating a cascade of events that leads to initiation of the DNA damage response. Current models favour a structure in which two replication forks converge on a single lesion, which is most likely to occur in late S phase[23]; however, it is likely that a replication fork approaching the lesion from only one side would also elicit repair. Replication-fork stalling at the site of damage is followed by an unhooking step, in which incisions are made on either side of the covalently linked nucleotides (Fig. 1). On the strand that is incised, TLS polymerases are recruited to bypass the unhooked crosslink. The incision event also effectively creates

Table 1 | Fanconi anaemia and Fanconi-anaemia-associated DNA repair factors and human disease

Genes	Human phenotypes	Cancer type in patients with Fanconi anaemia and in mutation carriers (italics)	Pathway and function	Link to Fanconi anaemia
*FANCA, FANCB, FANCC, FANCD2, FANCE, FANCF, FANCG, FANCI, FANCL** and *FANCM**	BMF, growth retardation, thumb and radial ray defects, developmental abnormalities, hypogonadism, microcephaly, infertility and hyperpigmentation	AML, HNSCC, oesophageal cancer, gynaecological cancers and liver cancers. *No cancers in mutation carriers for most groups, but FANCC carriers show increased risk of breast cancer*[75]	Fanconi anaemia; core complex; FANCL is an E3 ubiquitin ligase	Genes mutated in Fanconi anaemia
BRCA2 (also known as *FANCD1*)	BMF, growth retardation, thumb and radial ray defects and microcephaly	AML, ALL, medulloblastoma, neuroblastoma and Wilms tumour† *Breast, ovarian and prostate cancer*	Fanconi anaemia, homologous recombination; loads RAD51 onto DNA	Gene mutated in Fanconi anaemia
BRIP1 (also known as *FANCJ*)	BMF, growth retardation, thumb and radial ray defects, severe developmental abnormalities and hyperpigmentation	AML, HNSCC *Breast and ovarian cancer*	Fanconi anaemia, homologous recombination; 3′ to 5′ helicase	Gene mutated in Fanconi anaemia
PALB2 (also known as *FANCN*)	Anaemia, growth retardation, thumb radial ray defects, hyperpigmentation and microcephaly	AML, medulloblastoma, neuroblastoma, Wilms tumour and haemangioendothelioma† *Breast, ovarian and pancreatic cancer*	Fanconi anaemia, homologous recombination; promotes BRCA2 function	Gene mutated in Fanconi anaemia
RAD51C (provisionally known as *FANCO*‡)	Growth retardation, thumb and radial ray defects, hypogonadism, cystic kidneys and renal failure	*Breast and ovarian cancer*	Homologous recombination	Gene mutated in Fanconi-anaemia-like syndrome
SLX4 (also known as *FANCP*)	BMF, growth retardation, microcephaly, developmental defects including thumb and radial ray defects	HNSCC (one patient)	Fanconi anaemia; coordinates XPF–ERCC1, MUS81–EME1 and SLX1 nucleases	Gene mutated in Fanconi anaemia
BLM	Growth retardation, sun sensitivity, COPD, diabetes, recurrent infections and infertility	Broad predisposition including non-Hodgkin's lymphoma, AML, ALL, skin cancer, colon cancer and breast cancer	5′-3′ helicase, anti-recombinase	Interacts with core complex, BRIP1; genetic interaction with SLX4
BRCA1	No patients with biallelic mutations reported yet	*Breast and ovarian cancer*	Homologous recombination, interstrand crosslink repair	Interacts with BRIP1, BRCA2; necessary for Fanconi anaemia pathway activation
LIG4, DNA-PK$_{cs}$, Ku70 and Ku80§	Immune deficiency, pancytopaenias, growth retardation, microcephaly and skin photosensitivity	T-cell ALL (one report)	NHEJ	Genetic interaction with Fanconi anaemia proteins (see Table 2)
FAN1	KIN and renal failure	None reported	Interstrand crosslink repair	Interacts with I–D2 complex
XPF	Sun sensitivity, neurodegeneration and accelerated ageing	Basal and squamous cell carcinoma of the skin	NER, Fanconi anaemia nuclease	Interacts with SLX4

ALL, acute lymphocytic leukaemia; AML, acute myelogenous leukaemia; BMF, bone-marrow failure; COPD, chronic obstructive pulmonary disease; HNSCC, head and neck squamous cell carcinoma; I–D2, FANCI–FANCD2; KIN, karyomegalic interstitial nephritis; NER, nucleotide excision repair; NHEJ, non-homologous end joining. *No cancers have been reported in patients with *FANCL* or *FANCM* mutations. And the FANCM phenotype is unknown because the only patient described in the literature also has *FANCA* mutation. †In patients with *BRCA2* and *PALB2* mutations, the growth, bone-marrow and developmental phenotypes are as in other Fanconi anaemia complementation groups, but the probability of malignancy is higher, the spectrum is different and the age of onset is lower, indicative of a more extreme subtype of Fanconi anaemia. ‡FANCO is a provisional term, as patients with *RAD51C* mutations do not thus far display bone-marrow failure or cancer, and the chromosomal breakage in their cells is not as high as seen in individuals with Fanconi anaemia. §No mutations in *Ku70, Ku80* or *DNA-PK$_{cs}$* have ever been observed in patients.

a double-strand break, which is repaired by homologous recombination[23–25]. The multistep nature of interstrand crosslink repair requires the tight coordination of several different repair pathways, and it is within this delicately balanced system that the Fanconi anaemia pathway has its regulatory role.

An important step in the Fanconi anaemia pathway is the monoubiquitination of two components that form the I–D2 complex. Monoubiquitination of FANCD2 at Lys 561, and to a lesser extent of FANCI at Lys 523, represents the activation step of the Fanconi anaemia pathway[26,27]. This step is dependent on the Fanconi anaemia core complex, which comprises FANCA, FANCB, FANCC, FANCE, FANCF, FANCG, FANCL and FANCM, and accessory proteins, including FAAP20, FAAP24 and FAAP100. With the exception of FANCM, every member of the core complex is essential for successful monoubiquitination of I–D2. The ubiquitin ligase function of the core complex depends on the E3 ligase FANCL, with UBE2T lending its services as an E2 ligase[28–30]. Ubiquitination of FANCI and FANCD2 is mutually interdependent in mammalian cells, and also require discrete phosphorylation events: phosphorylation of FANCI at SQ or TQ sites close to the targeted lysine is necessary for specific ubiquitination of the I–D2 complex[31]. Recent evidence that FANCD2 ubiquitination is robustly stimulated by the presence of DNA suggests that this event may occur on chromatin[32]. Ubiquitinated I–D2 complex that is localized to chromatin orchestrates the actions of downstream repair proteins. Interestingly, the I–D2 complex may also have a direct role in the regulation of nucleosome assembly at sites of damage: the turnover of histone H3 after treatment with a crosslinker is markedly slowed after the gene encoding FANCD2 is knocked down[33].

The downstream components — those that are dispensable for monoubiquitination of the I–D2 complex — are BRCA2, BRIP1, PALB2 and RAD51C, which are the products of cancer predisposition genes necessary for homologous recombination, and SLX4 (also known as FANCP), which interacts with multiple nucleases (discussed in greater detail later). BRCA2 and PALB2 work in later stages of repair, in which they complete it through homologous recombination. The function of BRIP1 largely remains a mystery, and its role as a Fanconi anaemia protein needs to be explored further, although it has been implicated both in the activation of the Fanconi anaemia pathway as well as in the homologous recombination.

Although the details of how the Fanconi anaemia pathway is turned off have not been fully elucidated, one protein that clearly participates is the deubiquitinating enzyme USP1, which removes the ubiquitin from FANCD2 (ref. 34). In its absence, ubiquitinated FANCD2 is elevated, but the cells are sensitive to crosslink damage[35,36]. As such, it is clear that both appropriate activation and deactivation of the pathway are important for cell viability when challenged with interstrand crosslink damage.

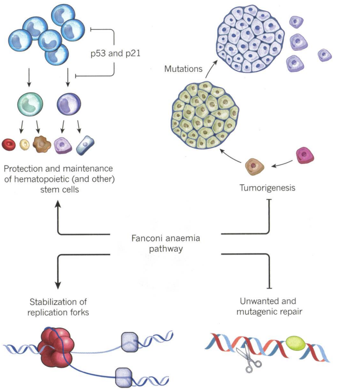

p53 and p21
Mutations
Protection and maintenance of hematopoietic (and other) stem cells
Tumorigenesis
Fanconi anaemia pathway
Stabilization of replication forks
Unwanted and mutagenic repair

Figure 2 | The diverse functions of the Fanconi anaemia pathway. The Fanconi anaemia pathway has many roles in human biology. Disrupting the Fanconi anaemia pathway leads to dysfunction of haematopoietic stem and progenitor cells as a result of elevated levels of p53 and p21 in stem cells, leading to impaired differentiation and a compromised pool of cells at birth. This could lead to stochastic developmental abnormalities and infertility in individuals with Fanconi anaemia. Several genes in the Fanconi anaemia pathway predispose individuals to some cancers. Fanconi anaemia proteins could have a role in suppressing tumorigenesis, the sequential accrual of mutations that transforms a healthy cell into a rapidly dividing cancer cell and finally into a metastatic tumour. The Fanconi anaemia pathway preserves replication-fork stability during S phase. The pathway also functions as a barrier against unwanted mutagenic repair processes, for example NHEJ.

The Fanconi anaemia pathway coordinates nuclease action

A key step in most repair pathways is the physical removal of the damaged bases, as well as the processing of the DNA by nucleases. In the case of interstrand crosslinks, the lesion affects both the Watson strand and the Crick strand, and its repair requires the collaboration of several processing enzymes. Each incision event requires a nuclease with the correct structure specificity to yield a repair intermediate that can correctly feed into the subsequent step. For example, crosslink unhooking must precede the resection steps that are necessary to provide a substrate for homologous recombination. Therefore, the removal steps are a complex choreography of nuclease action. The Fanconi anaemia pathway has been implicated in the recruitment and regulation of several different nucleases, including XPF–ERCC1, MUS81–EME1, SLX1 and FAN1 (refs 37–42).

SLX4 is a recently identified Fanconi anaemia protein that functions as a scaffold, modulator and cofactor for the three structure-specific nucleases XPF–ERCC1, MUS81–EME1 and SLX1 (refs 43–46). Essentially modular in nature, SLX4 is unique in that it can bring a veritable 'tool belt' of DNA-processing enzymes to the site of DNA damage, ensuring the presence of the best repair nuclease for the job. Of the three nucleases associated with SLX4, XPF–ERCC1 is the most important for resistance to interstrand crosslinks, with MUS81–EME1 and SLX1 having less prominent roles[47,48].

The exonuclease SNM1A is also biochemically active at sites of crosslinks. It is required for interstrand crosslink resistance, consistent with an *in vivo* function in repair of these lesions. Although its exact interaction with Fanconi anaemia proteins remains largely unclear, it is thought to act in concert with SLX4-associated XPF–ERCC1 to process the crosslink after the initial incision[49]. *In vitro*, SNM1A can digest interstrand crosslink containing DNA past the lesion, creating a preferred substrate for downstream TLS polymerases[49]. Understanding precisely how SNM1A, as well as other nucleases, is coordinated and controlled by, or independent of, Fanconi anaemia proteins will be crucial to understanding this multifaceted repair pathway.

Multiple research groups have identified the Fanconi-associated nuclease FAN1 to be required for interstrand crosslink resistance[38–41], implicating it in interstrand crosslink repair. FAN1 can be recruited to the sites of DNA damage in a manner dependent on the ubiquitinated I–D2 complex and on its own ubiquitin-binding domain. The precise mechanism of action of FAN1 on crosslinked DNA remains elusive; however, it is clear that the nuclease participates in a separate, but as yet uncharacterized, branch of crosslink repair in addition to the Fanconi anaemia pathway.

This function of FAN1 has recently been illuminated in the study of a rare genetic disease called karyomegalic interstitial nephritis (KIN), which is clinically characterized by progressive renal failure and histologically characterized by enlarged and hyperchromatic nuclei within the kidney tubular epithelium. A cohort of patients with KIN had mutations in *FAN1*. Interestingly, patients with KIN have none of the hallmarks of Fanconi anaemia, but their cells did have significant sensitivity to crosslinking agents. This sensitivity was complemented with wild-type FAN1, suggesting that an interstrand crosslink repair defect conferred by FAN1 deficiency does, at least in part, underlie the syndrome[50]. Deficiency of FAN1, which results in KIN but not Fanconi anaemia — despite the fact that FAN1 and FANCD2 interact — suggests that FAN1 might be redundant with another nuclease in the Fanconi anaemia pathway. It also implies that FAN1 has a function that is independent of the pathway in the kidney, in which FAN1 may repair lesions that the Fanconi anaemia pathway cannot. It is of interest that FAN1 expression is very high in the kidney. The deviation between KIN and Fanconi anaemia phenotypes is a fascinating puzzle that might help to understand different modes of interstrand-crosslink repair, as well as be a clue to the endogenous lesions that give rise to disease, not only in patients with rare disorders but also during normal ageing.

Fanconi anaemia pathway and replication stress

In addition to an indispensable role at DNA crosslinks, the Fanconi anaemia pathway is required for the protection of replication-fork stability under stress. The I–D2 complex is monoubiquitinated during S phase even in unchallenged cells[15]. The pathway is also activated by depletion of nucleotide pools with hydroxyurea, and even by ultraviolet light[51]. This is puzzling because the cell lines of patients with Fanconi anaemia are not sensitive to either of these agents. However, recent data suggest that FANCA, ubiquitinated FANCD2, BRCA1 and BRCA2 are necessary for the protection of the stalled replication fork after hydroxyurea treatment. It has been proposed that when these proteins are functionally absent, RAD51 is not deposited onto the fork under stress, and the nascent DNA is degraded in an MRE11-dependent manner[52,53]. The outcome seems to be an increase in chromosomal instability, which may not lead directly to immediate cellular demise, can still be mutagenic in the long term, thus contributing to the increase in tumorigenesis seen in patients with Fanconi anaemia. Importantly, the protective function of the Fanconi anaemia–BRCA axis, which has been uncovered by these experiments, seems to be separate from the canonical function of BRCA2 in homologous recombination. It will be important to address whether any of the interstrand crosslink sensitivity of Fanconi anaemia cell lines might be secondary to a defect in fork-protection function, and whether the stem-cell failure in Fanconi anaemia can be mitigated by fork stabilization.

Cooperation with other DNA repair pathways

At its heart, repair of interstrand crosslinks mediated by the Fanconi anaemia pathway is a highly regulated ensemble of other repair pathways: a nuclease borrowed from canonical nucleotide excision repair makes incisions, trans-lesion synthesis polymerases are responsible for filling in the gap opposite the unhooked crosslink, and homologous recombination is co-opted to resolve the double strand breaks that result from the encounter between replication forks and the lesion. It is becoming clear that — in addition to using other pathways for discrete repair steps — the Fanconi anaemia pathway is intimately entwined with the DNA damage response as a whole, engaging in collaboration and antagonism with almost all of the classic repair pathways in the fight for survival after genotoxic insult. The most notable of these are the mismatch repair pathway components implicated in FANCJ[54] and FANCD2 (ref. 55) function, and the helicase BLM, which seems to intersect with the Fanconi anaemia pathway at multiple levels. *BLM* mutations are found in Bloom syndrome, an autosomal recessive disease of growth retardation, immunodeficiency, sun sensitivity and profound cancer predisposition. The BLM protein is responsible for unwinding a variety of DNA structures and together with DNA topoisomerase IIIα is able to remove the double Holliday junctions that form during homologous recombination[56]. This process, called double Holliday junction dissolution, prevents the inappropriate engagement of nucleolytic Holliday junction resolvases, whose action may lead to deleterious exchanges of DNA between sister chromatids.

Table 2 | Summary of genetic interactions between Fanconi anaemia and non-homologous end joining

Fanconi anaemia gene	NHEJ gene	Outcome
Caenorhabditis elegans		
fcd-2 (*FANCD2*) (germline mutation)	*lig-4* (germline mutation)	Rescue of crosslink sensitivity and aberrant meiotic crossovers[63]
Chicken DT40 cells		
FANCC (knockout)	*Ku70* (knockout)	Partial rescue of crosslink sensitivity[64]
FANCC (knockout)	*LIG4* or *DNA-PK*$_{cs}$ (knockout)	No rescue[64]
Mouse embryonic fibroblasts		
Fanca$^{-/-}$, *Fancc*$^{-/-}$ (germline mutation)	*DNA-PK*$_{cs}$ (NU7026 inhibitor)	Partial rescue of crosslink sensitivity[63]
Fancd2$^{-/-}$ (germline mutation)	*Prkdc*$^{sc/sc}$ (*scid* mouse)	No rescue[65]
Fancd2$^{-/-}$ (germline mutation)	*Ku80*, 53bp1 (germline mutations)	No rescue. Increase in chromosomal instability and sensitivity[66]
Human cells		
M059J cells, *FANCD2* siRNA	*DNA-PK*$_{cs}$ deficient cell line	Rescue of crosslink (MMC) sensitivity[63]
HeLa cells, *FANCA* siRNA	*DNA-PK*$_{cs}$ (NU7026 inhibitor)	Partial rescue of crosslink (MMC) sensitivity[63]
HeLa cells, *FANCD2* siRNA	*DNA-PK*$_{cs}$ inhibitor	Rescue of crosslink sensitivity[63]
PD773 (*FANCD2*), PD331 (*FANCC*), patient mutations	*KU80* siRNA	Partial rescue of crosslink (MMC) sensitivity[63]

MMC, mitomycin C; siRNA, short interfering RNA.

Two connections between Fanconi anaemia proteins and BLM have recently been uncovered. One pertains to ultrafine anaphase bridges, which are thread-like DNA structures that form between masses of condensed DNA during anaphase and are thought to arise from persistent, unresolved DNA catenanes or recombination intermediates. Both BLM and the I–D2 complex localize to these microbridges, although FANCI and FANCD2 travel only to a subset of them — probably those caused by replication stress at fragile sites[57]. These microbridges may underlie the cytokinesis failure in Fanconi anaemia bone-marrow cells. The second connection is with SLX4, which collaborates with BLM during Holliday junction resolution. Co-depletion of these two proteins leads to a uniquely aberrant chromosome morphology, for which sister chromatids seem to stay linked and the chromosomes themselves adopt an odd, segmented, appearance[58]. As with many of the apparent interactions that the Fanconi anaemia pathway has with other repair pathways, the underlying mechanisms and relationships must still be elaborated, but taken together, the data reiterate how the Fanconi anaemia pathway has a far-reaching influence on general genome maintenance.

Fanconi anaemia and non-homologous end joining

The DNA repair pathway that has a particularly fraught relationship with the Fanconi anaemia pathway is NHEJ. This is an error-prone mechanism for fixing double-strand breaks for which no sister chromatid is available to be the template for homologous recombination, and DNA ends are ligated directly together without regard for homology. The heterodimer Ku70–Ku80 (Ku) binds to DNA ends, where it recruits the catalytic subunit of DNA-dependent protein kinase (DNA-PK$_{cs}$). DNA ligase-4 (LIG4) is then brought in to carry out the ligation (reviewed in ref. 59).

NHEJ is the default double-strand break repair pathway in the G0 and G1 phases. In other phases of the cell cycle, the cell faces a choice of whether to use NHEJ or homologous recombination to restore the break. Recent work has uncovered a dynamic shift between preferred pathways throughout the cell cycle, with homologous recombination reaching a peak during late S phase[60]. The decision also hinges on which factors access the free DNA end. The pattern of resection at the site of the double-strand DNA break provides a particular substrate for each type of repair. Homologous recombination requires extensive resection to create a long 3′ overhang that — once coated with RAD51 — can initiate homology search, although NHEJ requires little, if any, resection. Competition between the binding of homologous recombination and NHEJ factors influences the degree of resection at the break, and only once the 3′ overhang has been successfully resected is the cell fully committed to homologous recombination[61].

The Fanconi anaemia pathway is thought to have a role in pathway choice, funnelling the double-strand breaks created by interstrand crosslink processing or other forms of replication stress into high-fidelity homologous recombination repair. Therefore, NHEJ-mediated repair in the absence of a functional Fanconi anaemia pathway is speculated to be an underlying cause of the gross chromosomal abnormalities observed in Fanconi anaemia cell lines. Without a functional Fanconi anaemia pathway, perhaps the breaks become subject to the profligate end-pairing of NHEJ and thereby generate chromosomal rearrangements. Animal and human cell models have been developed to test this hypothesis, and so far the findings are contradictory (Table 2).

In the nematode *Caenorhabditis elegans*, the basic components of the Fanconi anaemia pathway are conserved, including the homologues *fcd-2* (*FANCD2*), *fnci-1* (*FANCI*), *fncm-1* (*FANCM*) and *dog-1* (*FANCJ*)[62]. The *C. elegans* I–D2 complex is ubiquitinated and travels to sites of damage similar to its human counterpart, and knockout of all four genes leads to hypersensitivity to crosslinking agents. *C. elegans* is an excellent model system to observe the effects of knocking out NHEJ components in a Fanconi-anaemia-deficient background, as it mimics relevant phenotypes with a pared-down pathway. When LIG4, the DNA ligase important for NHEJ, is mutated in *fcd-2* knockout worms, the sensitivity to crosslinking agents is corrected to near wild-type levels.

These data suggest that NHEJ does in fact have a deleterious effect when the Fanconi anaemia pathway is deficient[63].

The *C. elegans* model is also useful for studying the interplay between the Fanconi anaemia pathway and NHEJ during meiosis. Although patients with Fanconi anaemia, as well as knockout mice, exhibit fertility defects, the mechanisms underlying this loss of fertility are poorly understood. The regulation and resolution of meiotic crossovers derived from recombination of programmed double-strand breaks is mediated by many proteins that also serve in a DNA repair capacity, raising the possibility that the Fanconi anaemia pathway might also have a role. In the *fcd-2* mutant, RAD51 foci persist during meiosis, which are indicative of unresolved homologous recombination intermediates, and inappropriate non-homologous chromosome engagements are observed after diakinesis[63]. Similar to what is seen in somatic cells, the deletion of *lig-4* ameliorates the aberrant crossovers observed in the *fcd-2* mutant.

In cells of the chicken DT40 cell line, the interstrand crosslink hypersensitivity and chromosome breakage that are endemic in *FANCC* knockout cell lines can be suppressed by also knocking out *Ku70* (ref. 64). Interestingly, knocking out either *DNA-PK*$_{cs}$ or *LIG4* in the DT40 cells effects no such rescue, in contrast with the observations in *C. elegans*. In this case, the specificity of Ku70 rescue may point to a direct role for Fanconi anaemia proteins, possibly the I–D2 complex, in antagonizing Ku attachment at DNA ends.

The implication that NHEJ underlies the hypersensitivity and chromosomal breakage phenotypes in Fanconi anaemia is of great clinical interest because it points to a putative pharmaceutical target. If the cellular phenotypes that underlie the disease are caused not by interstrand crosslinks themselves but by aberrant repair processes that occur in the absence of the Fanconi anaemia pathway, perhaps NHEJ inhibition could serve as a highly specific treatment. It is therefore crucial to investigate the relationship between these two pathways in mammalian cell lines. RNA interference (RNAi) knockdown of *Ku80* in *FANCC*-deficient and *FANCD2*-deficient patient cell lines partially suppresses sensitivity to the killing effects of crosslinking agents. Furthermore, the addition of a DNA-PK$_{cs}$ inhibitor mitigates the hypersensitivity that results from knockdown of *FANCA* or *FANCD2* by RNAi in HeLa cells, or by knockout of *Fanca* or *Fancc* in mouse embryonic fibroblasts[63]. This again is in contrast to genetic experiments that show that the crosslink sensitivity of the *Fancd2* knockout is not changed in the setting of decreased DNA-PK$_{cs}$ activity (*Prkdc*$^{sc/sc}$)[65] and that loss of p53-binding protein 1 (53BP1) or Ku80 renders FANCD2-deficient cells even more sensitive to DNA-crosslinking agents than FANCD2 deficiency alone[66].

The inconsistencies between these studies raise questions about how to parse data from different organisms, genetic backgrounds and experimental set-ups. As hinted by these experiments, different components of the NHEJ pathway may have different roles in different organisms and even in different cells of the same organism; indeed, fibroblasts may not accurately reflect how cells that are more relevant to Fanconi anaemia, such as haematopoietic stem cells and progenitor cells, might behave. Furthermore, interactions between components of the two pathways may be exquisitely sensitive to the spatial and temporal relationships between them. When considering the possible therapeutic implications of these findings, it is important to remember that full inhibition of the Ku proteins is not feasible in human cells because they have an essential function at human telomeres[67]. Overall, well-controlled work, especially in human haematopoietic cells, will be required to understand the interplay of the Fanconi anaemia and NHEJ pathways. It will also be necessary to determine how these extremely toxic interstrand crosslinks are repaired if both pathways are compromised.

Alcohol metabolism and Fanconi anaemia deficiency

The average human is not exposed to the clastogenic agents that are used in the laboratory to model DNA repair defects. However, the Fanconi anaemia pathway is required for dealing with replication stresses that can arise during unchallenged cell-cycle progression, as shown by the activation of the pathway during normal S phase. The pathway also probably evolved to maintain genome integrity in the face of mutagenic agents that arise endogenously from processes of cellular metabolism. Oxidative stress within the cell can yield nucleophilic crosslinking agents such as malondialdehyde and acrolein, which derive from lipid peroxidation. Nitric oxide, another endogenous genotoxin that can promote the formation of crosslinks, is generated independently of lipids[68]. Aldehydes (discussed in greater detail later) are also a major source of endogenous crosslink damage. Understanding how damage from these endogenous agents feeds into the Fanconi anaemia pathway is crucial to understanding the disease.

Ethanol metabolism has long been known to produce reactive aldehydes that can function as carcinogens: acetaldehyde, an intermediate in the metabolic processing of ethanol, is particularly implicated as a genotoxic agent. Acetaldehyde stimulates monoubiquitination of FANCD2 *in vitro*, and Fanconi anaemia cells are sensitive to formaldehyde[69,70]. Until recently, however, no *in vivo* studies functionally linked these cellular by-products to concrete phenotypes in a Fanconi anaemia model. The first indication that such cellular by-products may indeed pose a challenge to Fanconi-anaemia-pathway-deficient systems came from a genetic interaction between Fanconi anaemia genes and superoxide dismutase, an enzyme that decreases the load of oxidative damage in the cell as a whole. *Fancc*$^{-/-}$ and *Sod*$^{-/-}$ mice display bone-marrow hypocellularity, which is not present in either of the single mutants[71]. Studies investigating the effects of interfering with acetaldehyde catabolism have uncovered a strong connection to the Fanconi anaemia pathway. Chicken DT40 cells with Fanconi anaemia genes knocked out, including *FANCB*, *FANCC*, *FANCL* and *BRCA2*, are sensitive to treatment with acetaldehyde[72]. In addition, DT40 cells with Fanconi anaemia knockouts are synthetic lethal with mutations in the formaldehyde catabolism gene *ADH5*, suggesting that toxicity is generated when the cells permit natural toxic aldehydes to build up[73].

A double mutant for *Fancd2* and the aldehyde dehydrogenase *Aldh2* was created to test the relevance of acetaldehyde sensitivity in the developing mouse. The double mutant did not survive, with embryos dying between embryonic day 9.5 and 13.5, but only if the mother was homozygous for the *Aldh2* mutation. By contrast, pregnant female mice with one wild-type *Aldh2* allele could carry pups to term. When heterozygous mothers were fed ethanol during pregnancy, the double-mutant pups exhibited drastic developmental abnormalities, such as exencephaly and eye defects[72]. These data suggest that the developing mouse is exquisitely sensitive to toxic aldehydes, casting light onto the significance of aldehyde catabolism in Fanconi-anaemia-pathway deficient backgrounds.

A more important question to ask is whether toxic aldehydes can produce symptoms in the mouse that mimic the human disease phenotype. Chronic alcohol abuse is correlated with bone-marrow problems, reminiscent of the bone-marrow failure observed in patients with Fanconi anaemia. Indeed, *Fancd2*$^{-/-}$:*Aldh2*$^{-/-}$ mice display increased levels of ethanol-induced bone-marrow failure compared with wild-type mice or single mutants. Even without administration of ethanol, the double mutants die early, owing to abnormalities that closely resemble acute lymphoblastic leukaemia.

Most tellingly, the double mutant mice that do not suffer from leukaemia develop aplastic anaemia that is similar to that seen in patients with Fanconi anaemia. The bone-marrow failure in these mice can be traced back to a DNA repair deficiency in the haematopoietic stem-cell and progenitor-cell pool, suggesting that endogenous aldehyde may be particularly harmful to haematopoietic stem cells, and providing the first direct link between acetaldehyde-related toxicity and bone-marrow failure phenotypes in Fanconi anaemia[74].

Animal, particularly mouse, models of Fanconi anaemia allow us to explore what genetic and environmental factors contribute to disease pathogenesis and, as such, can lead to better disease management and treatment plans. As sequencing becomes more affordable, understanding what genetic factors modify and shape the progress of Fanconi anaemia symptoms will offer better diagnostics and tailored treatment. Discovering what commonly encountered genotoxic agents contribute to the clinical manifestations of Fanconi anaemia will provide therapeutic targets and lifestyle modifications, not only for patients with Fanconi anaemia but also for those with many diseases that are derived from DNA repair defects.

Looking forward

As a multifunctional, highly connected pathway, the Fanconi anaemia pathway is well poised to be a very important topic in DNA repair and in biology in general. Although there is evidence of functional relationships between the Fanconi anaemia pathway and virtually all of the cellular DNA repair pathways, many of the details remain murky. Understanding the mechanisms underlying these connections will foster understanding of DNA repair as a whole, rather than as a mixed bag of disparate pathways. The genetic link between Fanconi anaemia and breast cancer will also be an important topic as research progresses, shedding light on the nature of tumour suppressors, and the enigma of genotype–phenotype correlations. As technologies make it easier to manipulate stem cells, the function of Fanconi anaemia in haematopoietic stem-cell lineages and other stem cells will become a growing subfield within Fanconi anaemia research. We may finally gain insight into the developmental abnormalities that are seen in patients. Combining studies of Fanconi anaemia function and stem-cell development will allow us to create a powerful model of DNA repair within stem-cell and progenitor-cell populations. Another pressing issue is the identification of the endogenous damage that the Fanconi anaemia pathway responds to. It is likely that the lesions arise from multiple sources and that there will be nuances to how the Fanconi anaemia pathway responds to each of them. Finally, there are many details about the molecular mechanism of interstrand crosslink repair that are still missing. Coordination and regulation of the nucleases identified as necessary for repair is a topic of intense interest. As new players in the pathway are identified, we can only expect that it will become more complex, but also more revealing. Deciphering the mechanics of Fanconi anaemia pathway function will help with efforts to understand both a fascinating disease, and basic human biology. ■

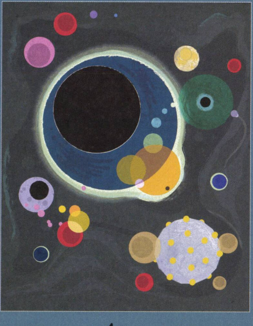
Received 30 July; accepted 8 November 2012.

1. Rosenberg, P. S., Tamary, H. & Alter, B. P. How high are carrier frequencies of rare recessive syndromes? Contemporary estimates for Fanconi anemia in the United States and Israel. *Am. J. Med. Genet. A* **155A,** 1877–1883 (2011).
2. Auerbach, A. D. Fanconi anemia and its diagnosis. *Mutat. Res.* **668,** 4–10 (2009).
3. Ceccaldi, R. *et al.* Bone marrow failure in Fanconi anemia is triggered by an exacerbated p53/p21 DNA damage response that impairs hematopoietic stem and progenitor cells. *Cell Stem Cell* **11,** 36–49 (2012).
This study identifies activation of the p53–p21 axis as a major contributing factor to loss of haematopoietic stem cells in Fanconi anaemia.
4. Tulpule, A. *et al.* Knockdown of Fanconi anemia genes in human embryonic stem cells reveals early developmental defects in the hematopoietic lineage. *Blood* **115,** 3453–3462 (2010).
5. Vinciguerra, P., Godinho, S. A., Parmar, K., Pellman, D. & D'Andrea, A. D. Cytokinesis failure occurs in Fanconi anemia pathway-deficient murine and human bone marrow hematopoietic cells. *J. Clin. Invest.* **120,** 3834–3842 (2010).
6. Muller, L. U. *et al.* Overcoming reprogramming resistance of Fanconi anemia cells. *Blood* **119,** 5449–5457 (2012).
7. Raya, A. *et al.* Disease-corrected haematopoietic progenitors from Fanconi anaemia induced pluripotent stem cells. *Nature* **460,** 53–59 (2009).
This paper demonstrates that human fibroblasts deficient in the Fanconi anaemia pathway are refractory to reprogramming into iPS cells unless they are first corrected with an appropriate gene, and that Fanconi anaemia-corrected cells can be reprogrammed into iPS cells that give rise to phenotypically normal lineages.
8. Deans, A. J. & West, S. C. DNA interstrand crosslink repair and cancer. *Nature Rev. Cancer* **11,** 467–480 (2011).
9. Meindl, A. *et al.* Germline mutations in breast and ovarian cancer pedigrees establish *RAD51C* as a human cancer susceptibility gene. *Nature Genet.* **42,** 410–414 (2010).
10. Rafnar, T. *et al.* Mutations in *BRIP1* confer high risk of ovarian cancer. *Nature Genet.* **43,** 1104–1107 (2011).
11. Rahman, N. *et al. PALB2*, which encodes a BRCA2-interacting protein, is a breast cancer susceptibility gene. *Nature Genet.* **39,** 165–167 (2007).
12. Seal, S. *et al.* Truncating mutations in the Fanconi anemia J gene *BRIP1* are low-penetrance breast cancer susceptibility alleles. *Nature Genet.* **38,** 1239–1241 (2006).
13. Tischkowitz, M. *et al.* Analysis of PALB2/FANCN-associated breast cancer families. *Proc. Natl Acad. Sci. USA* **104,** 6788–6793 (2007).
14. Walsh, T. *et al.* Mutations in 12 genes for inherited ovarian, fallopian tube, and peritoneal carcinoma identified by massively parallel sequencing. *Proc. Natl Acad. Sci. USA* **108,** 18032–18037 (2011).
15. Howlett, N. G. *et al.* Biallelic inactivation of BRCA2 in Fanconi anemia. *Science* **297,** 606–609 (2002).
This paper identifies the breast cancer predisposition gene *BRCA2* as mutated in a very severe form of Fanconi anaemia, firmly establishing Fanconi anaemia as a DNA repair deficiency disease.
16. Xia, B. *et al.* Fanconi anemia is associated with a defect in the BRCA2 partner PALB2. *Nature Genet.* **39,** 159–161 (2007).
17. Levran, O. *et al.* The BRCA1-interacting helicase BRIP1 is deficient in Fanconi anemia. *Nature Genet.* **37,** 931–933 (2005).
18. Levitus, M. *et al.* The DNA helicase BRIP1 is defective in Fanconi anemia complementation group J. *Nature Genet.* **37,** 934–935 (2005).
19. Vaz, F. *et al.* Mutation of the *RAD51C* gene in a Fanconi anemia-like disorder. *Nature Genet.* **42,** 406–409 (2010).
20. Enoiu, M., Jiricny, J. & Schäerer, O. D. Repair of cisplatin-induced DNA interstrand crosslinks by a replication-independent pathway involving transcription-coupled repair and translesion synthesis. *Nucleic Acids Res.* **40,** 8953–8964 (2012).
21. Wang, X. *et al.* Involvement of nucleotide excision repair in a recombination-independent and error-prone pathway of DNA interstrand cross-link repair. *Mol. Cell. Biol.* **21,** 713–720 (2001).
22. Shen, X. *et al.* Recruitment of Fanconi anemia and breast cancer proteins to DNA damage sites is differentially governed by replication. *Mol. Cell* **35,** 716–723 (2009).
23. Knipscheer, P. *et al.* The Fanconi anemia pathway promotes replication-dependent DNA interstrand cross-link repair. *Science* **326,** 1698–1701 (2009).
This work uses a cell-free system to demonstrate the requirement for FANCD2 and FANCI in the replication-dependent repair of interstrand crosslinks.
24. Raschle, M. *et al.* Mechanism of replication-coupled DNA interstrand crosslink repair. *Cell* **134,** 969–980 (2008).
25. Long, D. T., Raschle, M., Joukov, V. & Walter, J. C. Mechanism of RAD51-dependent DNA interstrand cross-link repair. *Science* **333,** 84–87 (2011).
26. Smogorzewska, A. *et al.* Identification of the FANCI protein, a monoubiquitinated FANCD2 paralog required for DNA repair. *Cell* **129,** 289–301 (2007).
27. Taniguchi, T. *et al.* S-phase-specific interaction of the Fanconi anemia protein, FANCD2, with BRCA1 and RAD51. *Blood* **100,** 2414–2420 (2002).
28. Alpi, A. F., Pace, P. E., Babu, M. M. & Patel, K. J. Mechanistic insight into site-restricted monoubiquitination of FANCD2 by Ube2t, FANCL, and FANCI. *Mol. Cell* **32,** 767–777 (2008).
29. Machida, Y. J. *et al.* UBE2T is the E2 in the Fanconi anemia pathway and undergoes negative autoregulation. *Mol. Cell* **23,** 589–596 (2006).
30. Meetei, A. R. *et al.* A novel ubiquitin ligase is deficient in Fanconi anemia. *Nature Genet.* **35,** 165–170 (2003).
31. Ishiai, M. *et al.* FANCI phosphorylation functions as a molecular switch to turn on the Fanconi anemia pathway. *Nature Struct. Mol. Biol.* **15,** 1138–1146 (2008).
32. Sato, K., Toda, K., Ishiai, M., Takata, M. & Kurumizaka, H. DNA robustly stimulates FANCD2 monoubiquitylation in the complex with FANCI. *Nucleic Acids Res.* **40,** 4553–4561 (2012).
33. Sato, K. *et al.* Histone chaperone activity of Fanconi anemia proteins, FANCD2 and FANCI, is required for DNA crosslink repair. *EMBO J.* **31,** 3524–3536 (2012).
34. Nijman, S. M. *et al.* The deubiquitinating enzyme USP1 regulates the Fanconi anemia pathway. *Mol. Cell* **17,** 331–339 (2005).
35. Kim, J. M. *et al.* Inactivation of murine Usp1 results in genomic instability and a Fanconi anemia phenotype. *Dev. Cell* **16,** 314–320 (2009).
36. Oestergaard, V. H. *et al.* Deubiquitination of FANCD2 is required for DNA crosslink repair. *Mol. Cell* **28,** 798–809 (2007).
37. Kim, Y. *et al.* Mutations of the *SLX4* gene in Fanconi anemia. *Nature Genet.* **43,** 142–146 (2011).
38. Kratz, K. *et al.* Deficiency of FANCD2-associated nuclease KIAA1018/FAN1 sensitizes cells to interstrand crosslinking agents. *Cell* **142,** 77–88 (2010).
39. Liu, T., Ghosal, G., Yuan, J., Chen, J. & Huang, J. FAN1 acts with FANCI–FANCD2 to promote DNA interstrand cross-link repair. *Science* **329,** 693–696 (2010).
40. MacKay, C. *et al.* Identification of KIAA1018/FAN1, a DNA repair nuclease recruited to DNA damage by monoubiquitinated FANCD2. *Cell* **142,** 65–76 (2010).
41. Smogorzewska, A. *et al.* A genetic screen identifies FAN1, a Fanconi anemia-associated nuclease necessary for DNA interstrand crosslink repair. *Mol. Cell* **39,** 36–47 (2010).
42. Stoepker, C. *et al.* SLX4, a coordinator of structure-specific endonucleases, is mutated in a new Fanconi anemia subtype. *Nature Genet.* **43,** 138–141 (2011).
43. Fekairi, S. *et al.* Human SLX4 is a Holliday junction resolvase subunit that binds multiple DNA repair/recombination endonucleases. *Cell* **138,** 78–89 (2009).
44. Svendsen, J. M. *et al.* Mammalian BTBD12/SLX4 assembles a Holliday junction resolvase and is required for DNA repair. *Cell* **138,** 63–77 (2009).
45. Andersen, S. L. *et al. Drosophila* MUS312 and the vertebrate ortholog BTBD12 interact with DNA structure-specific endonucleases in DNA repair and recombination. *Mol. Cell* **35,** 128–135 (2009).

46. Munoz, I. M. *et al.* Coordination of structure-specific nucleases by human SLX4/BTBD12 is required for DNA repair. *Mol. Cell* **35,** 116–127 (2009).
References 43–46 have identified SLX4 as a partner to multiple DNA repair nucleases necessary for interstrand crosslink repair.
47. Bhagwat, N. *et al.* XPF–ERCC1 participates in the Fanconi anemia pathway of cross-link repair. *Mol. Cell. Biol.* **29,** 6427–6437 (2009).
48. Kim, Y., Spitz, G. S., Veturi, U., Lach, F. P., Auerbach, A. D. & Smogorzewska, A. Regulation of multiple DNA repair pathways by the Fanconi anemia protein SLX4. *Blood* http://dx.doi.org/10.1182/blood-2012-07-441212 (23 October 2012).
49. Wang, A. T. *et al.* Human SNM1A and XPF–ERCC1 collaborate to initiate DNA interstrand cross-link repair. *Genes Dev.* **25,** 1859–1870 (2011).
50. Zhou, W. *et al.* *FAN1* mutations cause karyomegalic interstitial nephritis, linking chronic kidney failure to defective DNA damage repair. *Nature Genet.* **44,** 910–915 (2012).
51. Howlett, N. G., Taniguchi, T., Durkin, S. G., D'Andrea, A. D. & Glover, T. W. The Fanconi anemia pathway is required for the DNA replication stress response and for the regulation of common fragile site stability. *Hum. Mol. Genet.* **14,** 693–701 (2005).
52. Schlacher, K. *et al.* Double-strand break repair-independent role for BRCA2 in blocking stalled replication fork degradation by MRE11. *Cell* **145,** 529–542 (2011).
53. Schlacher, K., Wu, H. & Jasin, M. A distinct replication fork protection pathway connects Fanconi anemia tumor suppressors to RAD51-BRCA1/2. *Cancer Cell* **22,** 106–116 (2012).
This work identifies an important role for FANCD2, BRCA1 and BRCA2 in protecting the stability of replication forks, illuminating a repair-independent function for these proteins in genome maintenance.
54. Peng, M. *et al.* The FANCJ/MutLα interaction is required for correction of the cross-link response in FA-J cells. *EMBO J.* **26,** 3238–3249 (2007).
55. Williams, S. A. *et al.* Functional and physical interaction between the mismatch repair and FA-BRCA pathways. *Hum. Mol. Genet.* **20,** 4395–4410 (2011).
56. Wu, L. & Hickson, I. D. The Bloom's syndrome helicase suppresses crossing over during homologous recombination. *Nature* **426,** 870–874 (2003).
57. Chan, K. L., Palmai-Pallag, T., Ying, S. & Hickson, I. D. Replication stress induces sister-chromatid bridging at fragile site loci in mitosis. *Nature Cell Biol.* **11,** 753–760 (2009).
58. Wechsler, T., Newman, S. & West, S. C. Aberrant chromosome morphology in human cells defective for Holliday junction resolution. *Nature* **471,** 642–646 (2011).
59. Lieber, M. R. The mechanism of double-strand DNA break repair by the nonhomologous DNA end-joining pathway. *Annu. Rev. Biochem.* **79,** 181–211 (2010).
60. Karanam, K., Kafri, R., Loewer, A. & Lahav, G. Quantitative live cell imaging reveals a gradual shift between DNA repair mechanisms and a maximal use of HR in mid S phase. *Mol. Cell* **47,** 320–329 (2012).
61. Symington, L. S. & Gautier, J. Double-strand break end resection and repair pathway choice. *Annu. Rev. Genet.* **45,** 247–271 (2011).
62. Youds, J. L., Barber, L. J. & Boulton, S. J. *C. elegans*: a model of Fanconi anemia and ICL repair. *Mutat. Res.* **668,** 103–116 (2009).
63. Adamo, A. *et al.* Preventing nonhomologous end joining suppresses DNA repair defects of Fanconi anemia. *Mol. Cell* **39,** 25–35 (2010).
This work investigates the relationship between the Fanconi anaemia pathway and NHEJ, suggesting that the Fanconi anaemia cellular phenotypes derive from aberrant repair by NHEJ.
64. Pace, P. *et al.* Ku70 corrupts DNA repair in the absence of the Fanconi anemia pathway. *Science* **329,** 219–223 (2010).
65. Houghtaling, S. *et al.* Fancd2 functions in a double strand break repair pathway that is distinct from non-homologous end joining. *Hum. Mol. Genet.* **14,** 3027–3033 (2005).
66. Bunting, S. F. *et al.* BRCA1 functions independently of homologous recombination in DNA interstrand crosslink repair. *Mol. Cell* **46,** 125–135 (2012).
67. Wang, Y., Ghosh, G. & Hendrickson, E. A. Ku86 represses lethal telomere deletion events in human somatic cells. *Proc. Natl Acad. Sci. USA* **106,** 12430–12435 (2009).
68. Pang, Q. & Andreassen, P. R. Fanconi anemia proteins and endogenous stresses. *Mutat. Res.* **668,** 42–53 (2009).
69. Marietta, C., Thompson, L. H., Lamerdin, J. E. & Brooks, P. J. Acetaldehyde stimulates FANCD2 monoubiquitination, H2AX phosphorylation, and BRCA1 phosphorylation in human cells *in vitro*: implications for alcohol-related carcinogenesis. *Mutat. Res.* **664,** 77–83 (2009).
70. Ridpath, J. R. *et al.* Cells deficient in the FANC/BRCA pathway are hypersensitive to plasma levels of formaldehyde. *Cancer Res.* **67,** 11117–11122 (2007).
71. Hadjur, S. *et al.* Defective hematopoiesis and hepatic steatosis in mice with combined deficiencies of the genes encoding Fancc and Cu/Zn superoxide dismutase. *Blood* **98,** 1003–1011 (2001).
72. Langevin, F., Crossan, G. P., Rosado, I. V., Arends, M. J. & Patel, K. J. Fancd2 counteracts the toxic effects of naturally produced aldehydes in mice. *Nature* **475,** 53–58 (2011).
73. Rosado, I. V., Langevin, F., Crossan, G. P., Takata, M. & Patel, K. J. Formaldehyde catabolism is essential in cells deficient for the Fanconi anemia DNA-repair pathway. *Nature Struct. Mol. Biol.* **18,** 1432–1434 (2011).
74. Garaycoechea, J. I. *et al.* Genotoxic consequences of endogenous aldehydes on mouse haematopoietic stem cell function. *Nature* **489,** 571–575 (2012).
References 72 and 74 demonstrate that increasing aldehyde load in Fanconi anaemia mutant mice leads to phenotypes that are strikingly similar to those of the human disease, including developmental abnormalities, leukaemia and bone-marrow failure.
75. Berwick, M. *et al.* Genetic heterogeneity among Fanconi anemia heterozygotes and risk of cancer. *Cancer Res.* **67,** 9591–9596 (2007).

Acknowledgements We thank A. Auerbach, S. Gregg, E. Garner, Y. Kim, S. Thongthip and A. Wang for their comments. Work in our laboratory is supported by the Burroughs Wellcome Fund Career Award for Medical Scientists, the Starr Center Consortium grant, and by grant number 8 UL1 TR000043 from the National Center for Research Resources and the National Center for Advancing Translational Sciences (NCATS), National Institutes of Health. A.S. is a Rita Allen Foundation, Irma T. Hirschl, Alexandrine and Alexander Sinsheimer Foundation scholar, and is a recipient of the Doris Duke Clinical Scientist Development Award. M.C.K. is supported by an American Cancer Society-J.T.Tai postdoctoral fellowship.

Author Information Reprints and permission information is available at www.nature.com/reprints. The authors declare no competing financial interests. Readers are welcome to comment on the online version of this article at go.nature.com/vlcdeg. Correspondence should be addressed to A.S. (asmogorzewska@rockefeller.edu).



nature

www.nature.com/nature/supplements/insights

doi:10.1038/nature11770

Towards an exact description of electronic wavefunctions in real solids

George H. Booth[1], Andreas Grüneis[1,2], Georg Kresse[2] & Ali Alavi[1]

The properties of all materials arise largely from the quantum mechanics of their constituent electrons under the influence of the electric field of the nuclei. The solution of the underlying many-electron Schrödinger equation is a 'non-polynomial hard' problem, owing to the complex interplay of kinetic energy, electron–electron repulsion and the Pauli exclusion principle. The dominant computational method for describing such systems has been density functional theory. Quantum-chemical methods—based on an explicit ansatz for the many-electron wavefunctions and, hence, potentially more accurate—have not been fully explored in the solid state owing to their computational complexity, which ranges from strongly exponential to high-order polynomial in system size. Here we report the application of an exact technique, full configuration interaction quantum Monte Carlo to a variety of real solids, providing reference many-electron energies that are used to rigorously benchmark the standard hierarchy of quantum-chemical techniques, up to the 'gold standard' coupled-cluster ansatz, including single, double and perturbative triple particle-hole excitation operators. We show the errors in cohesive energies predicted by this method to be small, indicating the potential of this computationally polynomial scaling technique to tackle current solid-state problems.

Although density functional theory has been the workhorse of computational materials science for several decades[1], systematic routes to improve the crucial but approximate exchange-correlation functionals do not exist[2]. In contrast, for molecular systems, a systematic hierarchy of approximate yet highly successful quantum-chemical techniques, such as coupled-cluster theory, has long been established[3]. This hierarchy has not yet been explored in solids, although initial implementations of its lower levels have been encouraging[4–11]. Part of the reason for this is the high computational complexity of quantum-chemical methods. The computational cost grows rapidly with the number of considered electrons N and with the basis set size M. Traditional full configuration interaction (FCI), that is, exact diagonalization, has combinatorial scaling and can cope with at most some ten electrons in a small basis; and even coupled-cluster methods, although requiring only a computational time that is polynomial in M and N, are extremely expensive. However, recent developments in methodology, as well as the increase in computer power, mean that it is now possible to address their accuracy and applicability in this domain.

In the development of quantum chemistry, FCI has played an invaluable benchmarking role, by providing exact results within a given basis. This has enabled the electron correlation problem to be addressed in isolation from other complicating factors inherent when comparing to experiment[12]. Moreover, FCI enables us to assess the degree to which the electronic wavefunction is dominated by a single determinant, and therefore which systems are likely to be well described by approximations such as many-body perturbation theory and coupled-cluster theory[13]. In solids, however, the absence of FCI means that the accuracy of such approximations cannot be easily gauged, especially in systems where correlations are expected to be strong. Here we provide, for the first time to our knowledge, FCI-quality energies in a range of realistic solids and unambiguously evaluate the accuracy of high-level quantum-chemical methods.

A further motivation for the application of quantum-chemical methods to solids comes from the multitude of recent developments, which hold the promise to reduce the computational cost beyond our present consideration. These include optimized virtual spaces[14], explicit correlation[15,16], exploitation of locality of correlation[5], and others[17–19], and should be directly transferable to the solid state. Combining high accuracy with increased efficiency, quantum-chemistry methods hold the promise to routinely bring high accuracy to computational materials science.

FCIQMC and the quantum-chemical hierarchy

The FCI quantum Monte Carlo (FCIQMC) method has emerged as a tool to calculate energies that are essentially identical to the true correlation energy captured by the basis set, whilst having a significantly lower computational scaling than a traditional brute force diagonalization of the problem (FCI). This makes it ideally suited for a systematic benchmarking of wavefunction-based methods in the solid state[20–25]. This method involves a stochastic sampling of a Slater determinant space constructed from the basis set—a function space of orthonormal antisymmetrized determinants in which the wavefunctions are expressed. This method has previously been applied to molecular systems[20–24] and the homogeneous electron gas[25], where energies were calculated that compared favourably, or in some cases surpassed in accuracy, those achieved with state-of-the-art diffusion Monte Carlo (DMC) techniques[26]. This provides the confidence to tackle more realistic solid-state systems here.

Although it is possible for DMC to be used as a benchmark for quantum-chemistry methods and vice versa[27], DMC does not operate in a Slater determinant space, but rather a real space representation of the wavefunction. As such, it would require quantum-chemical calculations to be converged to high accuracy with respect to the basis set size before any meaningful comparisons could be drawn. Instead, by comparing to values obtained in the same Slater determinant basis, robust comparisons between FCIQMC and more approximate diagrammatic methods can be drawn without the need for absolute convergence. More approximate methods can then also be used to

---

[1]University of Cambridge, Chemistry Department, Lensfield Road, Cambridge CB2 1EW, UK. [2]University of Vienna, Faculty of Physics and Center for Computational Materials Science, Sensengasse 8/12, A-1090 Vienna, Austria.

extrapolate to the infinite basis set limit[14,17,28]. Additionally, DMC requires an approximation for the nodal surface of the wavefunction. Although this error can be made relatively small[29,30], releasing the nodal surface is notoriously difficult for solids and greatly increases the computational demand. Similarly, another variant of QMC methods, auxiliary-field quantum Monte Carlo (AFQMC), although now operating in a space of Slater determinants and with favourable scaling, requires analogous constraints within the phaseless approximation in order to go to realistic system sizes and avoid transient energy estimates[31,32].

Here we illustrate the extension of the FCIQMC method to the solid state. The introduction of translational and crystal momentum symmetries, which arise from working with finite simulation cells with periodic boundary conditions, necessitates a change to the 'walker' dynamics and 'initiator' rules (see below for descriptions). In order to take advantage of these properties, the method is reformulated for complex wavefunctions. The presence of a 'phase' problem, rather than a simpler sign problem, is considered. A range of systems, from the extensively studied lithium hydride, to other ionic, covalent and rare gas solids is considered, while benchmarking the established quantum-chemical methods of second-order Møller–Plesset theory (MP2)[33], coupled-cluster singles and doubles (CCSD)[34], and the first implementation of perturbative triples (CCSD(T))[35] for periodic systems. These three methods are generally considered to possess a favourable trade-off between accuracy and cost, and are the most widely used of all quantum-chemical methods. Finally, we study the far more complex electronic structure of the charge-transfer solid NiO, to evaluate the ability of these quantum-chemical methods to handle strong correlation effects, where deficiencies are likely to be exposed.

Sampling in the solid state

The recently developed FCIQMC method[20–24] involves a discrete sampling of the wavefunction by signed 'walkers' which stochastically evolve within a Hilbert space of N-electron Slater determinants as illustrated in Fig. 1. The method converges rapidly with the number of walkers to the FCI limit, and generally the number of walkers required is a tiny fraction of the Hilbert space. This leads to a huge compression in the wavefunction information, while correctly reproducing exact time-averaged properties. The reason that a sufficient number of walkers is required has to do with overcoming the 'fermion sign problem' present in the Monte Carlo sampling of any fermionic wavefunction[36]. The sign problem in this space manifests itself in the presence of a lower-energy state characterized by a combination of the $\pm\Psi$ degenerate solutions, which if not suppressed, leads to an exponential increase in noise. The growth of this state is controlled in FCIQMC by annihilation between walkers of opposite signs, which stabilizes the wavefunction to one signed solution[20,37]. The discrete space of Slater determinants allows FCIQMC to implement this annihilation exactly, and provided the walker population is large enough, will directly overcome the sign problem without requiring constraints on the wavefunction.

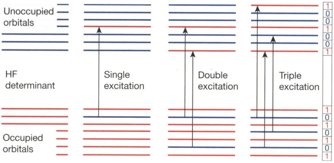


Figure 1 | FCIQMC in a nutshell. The antisymmetric many-electron wavefunction $\Psi(\mathbf{r}_1, \mathbf{r}_2, \ldots, \mathbf{r}_N)$ is represented in a space of determinants constructed out of Hartree–Fock (HF) orbitals (occupied as well as unoccupied one electron wavefunctions, $\phi(\mathbf{r})$). Walkers are encoded by a bit string as shown on the right, where each bit corresponds to one orbital. '1' (red) and '0' (black) imply that the orbital is occupied and unoccupied, respectively. Because some Slater determinants (for instance, the Hartree–Fock determinant) have a very high occupation probability, additional bits are reserved to count the signed number of real and imaginary walkers on a determinant. The computational procedure involves selecting a new determinant from an existing walker with a certain sign and transition probability as discussed in the main text. The excitation number of a determinant refers to the numbers of holes which need to be introduced in the Hartree–Fock determinant to generate the determinant.

Because walker annihilation can only occur on occupied determinants, it is important to ensure that the newly occupied space remains sign-coherent to the currently sampled wavefunction. This is the rationale behind the 'initiator' rules used in the i-FCIQMC method[22], which is used exclusively in this work, whereby newly occupied determinants must have originated from a determinant with a population greater than a parameter n_{add}. By restricting the growth of the occupied space in this way, the walker density and hence annihilation rate is kept high, ensuring that propagation of noise in the system is kept to a minimum. This biases the dynamic in a small way, but rigorously converges onto exact energies of the Hamiltonian as the walker number increases.

The determinants in this work are composed from antisymmetrized products of one-electron orbitals obtained from a prior Hartree–Fock calculation in a large basis of periodic plane waves within the framework of the projector-augmented wave method, as implemented in VASP. If these orbitals are strictly real then the wavefunction Ψ can also be real[20]. In this work, however, the orbitals are complex Bloch functions, to account for the translational invariance of the potential. With these we can construct many body wavefunctions and use k-point sampling to ensure convergence, rather than sampling ever larger unit cells to remove finite-size effects.

Because it is necessary to correlate between sampled k-points, the number of explicitly correlated electrons and orbitals increases linearly with the number of sampled k-points yielding combinatorial scaling in the size of the Hilbert space (essentially exponential with the number of k-points). However, performing this sampling increases the number of zero Hamiltonian matrix elements between determinants, as crystal momentum must be conserved. By implementing an algorithm to stochastically generate only these momentum-allowed excitations, a saving that grows quadratically with the number of k-points is achieved, because both the accessible space is reduced, and the magnitude of the time step is increased.

To take advantage of these savings, it is necessary to work with complex orbitals, requiring a complementary set of both 'real' and 'imaginary' walkers in the FCIQMC dynamic, and a reformulation of the algorithm. The master equations of the FCIQMC method follow naturally from the imaginary-time Schrödinger equation, and are given by

$$-\frac{\mathrm{d}N_{\mathbf{i}}}{\mathrm{d}\tau} = (H_{\mathbf{ii}} - S)N_{\mathbf{i}} + \sum_{\mathbf{j}\neq\mathbf{i}} H_{\mathbf{ij}}N_{\mathbf{j}} \tag{1}$$

where $N_{\mathbf{i}}$ represents the now complex walker population on determinant $D_{\mathbf{i}}$, τ represents imaginary time, S is a strictly real energy-offset parameter denoted the 'shift', which controls population growth, and $H_{\mathbf{ij}} = \langle D_{\mathbf{i}}|\hat{H}|D_{\mathbf{j}}\rangle$ is the many-electron Hamiltonian evaluated between two determinants. In each iteration, for each walker (real and imaginary) on a determinant $D_{\mathbf{i}}$, a suitable momentum-allowed excitation, $D_{\mathbf{j}}$, is generated. The real ($\Re$) and imaginary ($\Im$) parts of $H_{\mathbf{ij}}$ are considered in turn, and two attempts at generating new walkers on $D_{\mathbf{j}}$ are stochastically realized. For real parent walkers:

$$p_{\mathrm{s}}^{\Re}(\mathbf{j}|\mathbf{i}) = \frac{\delta\tau\left|\Re\left[H_{\mathbf{ij}}\right]\right|}{p_{\mathrm{gen}}(\mathbf{j}|\mathbf{i})}; \quad \mathrm{sign} = -\,\mathrm{sign}\left(\Re[N_{\mathbf{i}}]\Re\left[H_{\mathbf{ij}}\right]\right) \tag{2}$$

$$p_{\rm s}^{\Im}(\mathbf{j}|\mathbf{i}) = \frac{\delta\tau\left|\Im\left[H_{\mathbf{ij}}\right]\right|}{p_{\rm gen}(\mathbf{j}|\mathbf{i})}; \quad \text{sign} = -\,\text{sign}\left(\Re[N_{\mathbf{i}}]\Im\left[H_{\mathbf{ij}}\right]\right) \qquad (3)$$

and for imaginary parent walkers:

$$p_{\rm s}^{\Re}(\mathbf{j}|\mathbf{i}) = \frac{\delta\tau\left|\Im\left[H_{\mathbf{ij}}\right]\right|}{p_{\rm gen}(\mathbf{j}|\mathbf{i})}; \quad \text{sign} = -\,\text{sign}\left(\Im[N_{\mathbf{i}}]\Im\left[H_{\mathbf{ij}}\right]\right) \qquad (4)$$

$$p_{\rm s}^{\Im}(\mathbf{j}|\mathbf{i}) = \frac{\delta\tau\left|\Re\left[H_{\mathbf{ij}}\right]\right|}{p_{\rm gen}(\mathbf{j}|\mathbf{i})}; \quad \text{sign} = -\,\text{sign}\left(\Im[N_{\mathbf{i}}]\Re\left[H_{\mathbf{ij}}\right]\right) \qquad (5)$$

where τ is the timestep for the simulation, and $p_{\rm s}^{\Re}$ and $p_{\rm s}^{\Im}$ indicate the probability of creating real and imaginary child walkers, respectively. $p_{\rm gen}(\mathbf{j}|\mathbf{i})$ is the probability of generating determinant $D_{\mathbf{j}}$ from $D_{\mathbf{i}}$. After this step, the 'death' step is performed for each occupied determinant, with the same death probability for the real and imaginary walkers of $\delta\tau(H_{\mathbf{ii}} - E_0 - S)$ stochastically realized, where E_0 is a reference energy, and $H_{\mathbf{ii}}$ is now strictly real. A final annihilation step occurs every iteration, where real and imaginary walkers are separately considered, and pairs of opposite sign on the same determinant are removed from the simulation.

The value of S can be used as a strictly real measure of the correlation energy of the problem; however, provided a good overlap of the walker distribution with a reference wavefunction D_0 (generally taken to be the Hartree–Fock determinant) can be found, an averaged projected estimator is often less noisy:

$$E(\tau) = \frac{\langle D_0|\hat{H}|\Psi(\tau)\rangle}{\langle D_0|\Psi(\tau)\rangle} \qquad (6)$$

where $\langle D_0|\Psi(\tau)\rangle = N_0$. As opposed to S, $E(\tau)$ is now a complex quantity, where in order to achieve real energies, the imaginary part of the energy must cancel to zero in a non-trivial way. In order to test this, we considered rock-salt-structured LiH sampled using $2\times2\times2$ k-points. By choosing all k-points to lie at the Γ-point or Brillouin zone boundary, it was possible to take linear combinations of the orbitals to give a strictly real basis. This is compared to the complex basis in Fig. 2.

It can be seen that all methods, including FCIQMC, agree exactly between the two bases, and that in the complex basis, the imaginary component of the energy converges to zero within small error bars. Although there is the potential for rotations of the wavefunction in the complex plane, it is observed that the discretization of the wavefunction amplitudes prevents this from happening, and global U(1) transformations are thus suppressed after an initial arbitrary phase factor is determined. This indicates that there is no more of a sign-issue to overcome with annihilation events than that of the original real formulation of the dynamics.

Real solids and the accuracy of quantum chemistry

Having established the accuracy and efficiency of the complex FCIQMC walker dynamics, it was initially tested on the most widely studied solid to date, rock-salt-structured LiH (refs 6–9, 11, 17, 27). We first benchmark the accuracy of the MP2, CCSD and CCSD(T) energies for this system, by considering the deviation of each from FCIQMC values. Figure 3 shows an equation of state for a range of volumes, with a $3\times3\times3$ k-point sampling, in a minimal basis required to capture any non-dynamic correlation.

The efficient sampling of the i-FCIQMC method is clear, where the $3\times3\times3$ k-point mesh correlates 54 electrons in a space of $\sim10^{30}$ determinants, with convergence to the exact energy obtained after only ~50 million walkers, as demonstrated in Fig. 4a. The MP2 values are clearly shifted to higher energies compared to FCIQMC, and because this error changes significantly with volume, the MP2 equilibrium volume and bulk modulus deviate by 3.5% and 6.5%, respectively, from the FCIQMC values.

CCSD energies are virtually parallel to the FCIQMC results, yielding a similar volume and bulk modulus as FCIQMC. Finally, CCSD(T) shows almost exact agreement with FCIQMC in the absolute energies. Overall, these results very much mimic the performance of the standard quantum-chemical hierarchy established for molecular systems and are indicative of their suitability for other similar solids.

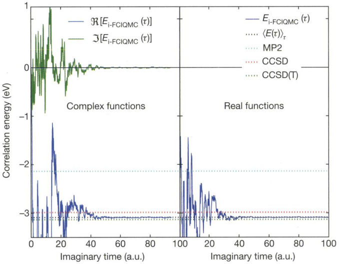


Figure 2 | Comparison of real and complex i-FCIQMC dynamics for the correlation energy of LiH. A $2\times2\times2$ Γ-centred k-point mesh with 16 electrons, and 40 correlated Hartree–Fock orbitals was employed, at a primitive rock-salt unit cell volume of 17.03 Å^3. Converged energies for 30 million walkers between the two bases agree within small stochastic error bars. The additional overhead for the complex dynamic means that the cost was ~5 times that of the real dynamic to converge to equivalent error bars. Also included are MP2, CCSD and CCSD(T) results for comparison. $\langle E(\tau)\rangle_\tau$ is an imaginary-time average of the projected energy $E(\tau)$, taken after a period of equilibration.

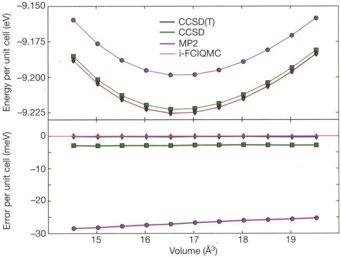


Figure 3 | Equation of state for rock-salt LiH, with the error from the i-FCIQMC value per unit cell at each volume shown below, for MP2, CCSD and CCSD(T). A $3\times3\times3$ Γ-centred k-point mesh was employed, with 54 electrons in 54 Hartree–Fock orbitals. All i-FCIQMC energies were converged with 55 million walkers, and error bars are too small to be seen on the plot ($\mathcal{O}$[0.1 meV]).

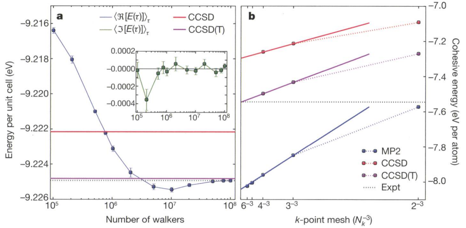


Figure 4 | Convergence with respect to the total walker number and *k*-point meshes. **a,** Convergence of the i-FCIQMC energy with total walker number of a $3 \times 3 \times 3$ *k*-point sampled LiH calculation. This explicitly correlated 54 electrons in 54 orbitals. Error bars (1s.d.) were calculated via a Flyvbjerg-Petersen 'blocking' algorithm[43]. The inset shows the imaginary component of the energy on a finer scale. **b,** Convergence of the cohesive energy of diamond with respect to the *k*-point mesh using MP2, CCSD and CCSD(T). The solid lines extrapolate to an infinite *k*-points mesh using $3 \times 3 \times 3$ and $4 \times 4 \times 4$ *k*-points. MP2 results for $5 \times 5 \times 5$ and $6 \times 6 \times 6$ *k*-points confirm that the finite size error decays as $1/N_k^3$, where N_k is the number of *k*-points in each direction.

To cover different bonding situations, in Fig. 5 we consider the relative errors of quantum-chemical methods for the correlation energies of several crystals when compared to FCIQMC. MP2 and CCSD recover between 80% and 98% of the FCIQMC correlation energy for the sample of rare gas, covalent and ionic solids. The dependence of the relative errors on the various systems is most pronounced in the case of MP2 theory. This is not unexpected and reflects the limitations of low-order perturbation theory. MP2 is more accurate for wide-gap insulators, such as Ne and LiF, than for semiconductors with a smaller gap, like Si and AlP. In contrast, CCSD(T) is shown to give a balanced description across the different systems and is in error by at most 2%.

What accuracy, therefore, can one expect from converged CCSD(T) calculations for solids? To answer this question, we computed cohesive energies. These are extremely demanding quantities for any theory, because the correlation effects in solids differ markedly from atoms, potentially leading to large errors in the prediction of the cohesive energy. Furthermore, attention needs to be paid to finite-size scaling (that is, *k*-point sampling). For the present study we have limited our attention to four solids, rock-salt LiH, diamond, zinc-blende BN and AlP, expecting similar results for the other materials. We have used the progressive downsampling technique[14,28], employing *k*-point meshes of up to $4 \times 4 \times 4$. Figure 4b shows for diamond that the MP2 cohesive energy converges as $1/N_k^3$, where N_k is the number of *k*-points used to sample the Brillouin zone in each direction. By fitting to the MP2, CCSD and CCSD(T) energies for $3 \times 3 \times 3$ and $4 \times 4 \times 4$ *k*-points, we can extrapolate to infinitely dense *k*-point meshes. The remaining finite size error on the correlation energy is expected to be less than 20 meV per atom for the considered systems. The computational cost is of the order of 25,000 CPU (central processing unit) hours for diamond, with results reported in Table 1.

As anticipated (owing to the established agreement with i-FCIQMC in smaller bases and supercells), the CCSD(T) results are all in almost exact agreement with the experimental cohesive energy corrected for the zero-point energy[38]. The MP2 cohesive energies generally show substantial error compared to experiment[5,9,17]. MP2 severely underestimates the correlation energy of atoms, and while it also underestimates the correlation energy of the solids as shown in Fig. 5, the perturbative nature of the theory leads to less underestimation of the correlation energy in solids so that the cohesive energy is often—but not always—overestimated compared to the experimental value. This makes MP2 an unreliable method for solids, especially for calculation of cohesive energies. CCSD and CCSD(T), on the other hand, are far more consistent. Although the absolute errors in the CCSD cohesive energies are still quite sizeable, the cohesive energies are always underestimated. This results from the CCSD correlation energy in the solids being always too small (Fig. 5) whereas in atoms it is generally a good approximation. On adding the perturbative triples correction to the CCSD, the cohesive energies dramatically improve, to an error of only 0.03 eV. As for the absolute correlation energies in solids, the (T) correction over-compensates, leading to too-negative correlation energies, and hence an overestimation of the cohesive energies, albeit only mildly so. We finally note that our results indicate that the residual errors are dominated by correlation errors in the

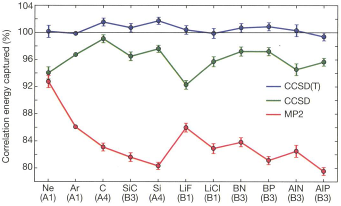


Figure 5 | Relative errors compared to i-FCIQMC in a range of solids. $2 \times 2 \times 2$ *k*-point sampling was employed, correlating 64 electrons in 64 orbitals, and 80 electrons in 72 orbitals, for LiF and LiCl, respectively. Lattice structures are given as A1 = f.c.c., A4 = diamond, B1 = rock salt, B3 = zinc blende, while lattice constants can be found in ref. 44. Error bars of each method are derived from the random errors of the i-FCIQMC values they are compared to. The i-FCIQMC calculations employed ~10,000 walkers on the Hartree–Fock determinant resulting in between 2 million and 300 million total walkers, depending on the system.

Table 1 | Cohesive energies evaluated within the quantum-chemical hierarchy

	LiH	C	BN	AlP
MP2	−2.386	−8.039	−7.149	−4.629
CCSD	−2.454	−7.295	−6.572	−4.107
CCSD(T)	−2.483	−7.545	−6.782	−4.347
Experiment	−2.487	−7.545	−6.758	−4.322

Energies are given in eV per atom. Experimental values have been corrected for zero-point vibration energies.

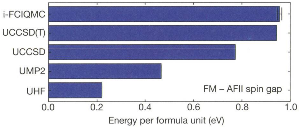


Figure 6 | Energies of the NiO spin gap, calculated at the level of UHF, UMP2, UCCSD, UCCSD(T) and i-FCIQMC in an UHF basis for both spin species. Correlating the 3*d* and 4*s* of the Ni, and 2*s* and 2*p* of the O using a rhombohedral supercell with a lattice constant of 4.19 Å containing 2 formula units, resulted in a sampling of 32 electrons, in 38 Hartree–Fock orbitals. 300 million walkers were required for the i-FCIQMC calculation of the AFII state, requiring ~50,000 CPU hours, while only 100 million walkers were needed for the FM state, with random error bar (1 s.d.) shown, and remaining error estimated to be less than 30 meV.

solid, whereas correlation energies for atoms are essentially converged at the CCSD(T) level.

The conclusion is that the accuracy of CCSD(T) has been established for solids to be of the order of 0.03 eV or 1 kcal mol^{-1}. By contrast, the most widely used density functional (PBE[39]) exhibits a mean-absolute error of 0.15–0.2 eV for a similar range of insulating solids[40].

Towards strong correlation

An important and open question is the domain of applicability of CCSD and CCSD(T) as stronger correlation effects set in. An initial indication of this can be found by comparing the performance of the methods for the spin gap between the ground antiferromagnetic state (AFII) and the ferromagnetic state (FM) of nickel oxide in a rhombohedral unit cell. This is a classic charge-transfer insulator, and it is expected to have strong correlation effects for at least one of the states, and as such, a balanced description for the calculation of the spin gap is expected to provide a stern test.

The results for the FM–AFII spin gap are given in Fig. 6 and illustrate the systematic convergence of the quantum-chemical hierarchy with respect to the correlation treatment of the system. Analysis of the FCIQMC wavefunction indicates that the ground state (AFII) is more strongly correlated and multiconfigurational, with a normalized Hartree–Fock weight of only 0.69, compared to 0.86 for the FM state. This lack of a dominant single reference leads to errors in MP2 of over 50%, while CCSD is still 19% in error compared to i-FCIQMC. Despite this, the qualitative behaviour of the quantum chemical hierarchy remains intact, with CCSD(T) providing an excellent approximation to the exact result. Results from unrestricted Hartree–Fock (UHF) calculations of NiO converged to the thermodynamic limit only capture 17% of the experimental spin gap between the two states[41], inferred from neutron scattering[42]. This agrees well with our restricted k-point sampled system, where the UHF captures 23% of the spin gap compared to FCIQMC results.

Conclusions and outlook

We have shown that FCI-quality correlation energies can be obtained for solid-state systems using an extension of the i-FCIQMC method to complex wavefunctions. We have demonstrated that the standard quantum-chemical hierarchy of increasingly accurate polynomially scaling methods holds for a range of materials, including rare gas, ionic and covalent solids and the charge-transfer insulator NiO. As explicitly shown by the cohesive energies of LiH, C, AlP and BN, CCSD(T) is very accurate for the solid state, surpassing 1 kcal mol^{-1} accuracy in reproducing experimental results. Considering the proven reliability of CCSD(T) for molecules, we expect a similar precision for insulators and semiconductors in general, with metals possibly requiring further methodological improvements. In combination with recent developments to reduce the computational cost—for instance with explicit inclusion of the cusp condition for the many-electron wavefunction—as well as further technical, algorithmic and methodological advances, the accuracy of FCIQMC and the quantum-chemistry methods will be brought routinely to solid-state physics and computational materials science. We are witnessing a slow but steady change of our computational paradigm.

Received 17 July; accepted 7 November 2012.

Published online 19 December 2012.

1. Kohn, W. Nobel lecture: Electronic structure of matter–wave functions and density functionals. *Rev. Mod. Phys.* **71,** 1253–1266 (1999).
2. Cohen, A. J., Mori-Sánchez, P. & Yang, W. Challenges for density functional theory. *Chem. Rev.* **112,** 289–320 (2012).
3. Pople, J. A. Nobel lecture: Quantum chemical models. *Rev. Mod. Phys.* **71,** 1267–1274 (1999).
4. Müller, C. & Paulus, B. Wavefunction-based electron correlation methods for solids. *Phys. Chem. Chem. Phys.* **14,** 7605–7614 (2012).
5. Maschio, L. *et al.* Fast local-MP2 method with density-fitting for crystals. I. Theory and algorithms. *Phys. Rev. B* **76,** 075101 (2007).
6. Casassa, S., Halo, M., Maschio, L., Roetti, C. & Pisani, C. Beyond a Hartree-Fock description of crystalline solids: the case of lithium hydride. *Theor. Chem. Acc.* **117,** 781–791 (2007).
7. Ayala, P., Kudin, K. & Scuseria, G. Atomic orbital Laplace-transformed second-order Møller-Plesset theory for periodic systems. *J. Chem. Phys.* **115,** 9698–9707 (2001).
8. Usvyat, D. *et al.* Approaching the theoretical limit in periodic local MP2 calculations with atomic-orbital basis sets: the case of LiH. *J. Chem. Phys.* **134,** 214105 (2011).
9. Marsman, M., Grüneis, A., Paier, J. & Kresse, G. Second-order Møller–Plesset perturbation theory applied to extended systems. I. Within the projector-augmented-wave formalism using a plane wave basis set. *J. Chem. Phys.* **130,** 184103 (2009).
10. Stoll, H. & Doll, K. Approaching the bulk limit with finite cluster calculations using local increments: the case of LiH. *J. Chem. Phys.* **136,** 074106 (2012).
11. Nolan, S. J., Gillan, M. J., Alfè, D., Allan, N. L. & Manby, F. R. Calculation of properties of crystalline lithium hydride using correlated wave function theory. *Phys. Rev. B* **80,** 165109 (2009).
12. Knowles, P. J. & Handy, N. C. A new determinant-based full configuration-interaction method. Chem. *Phys. Lett.* **111,** 315–321 (1984).
13. Larsen, H., Olsen, J., Jørgensen, P. & Christiansen, O. Full configuration interaction benchmarking of coupled-cluster models for the lowest singlet energy surfaces of N_2. *J. Chem. Phys.* **113,** 6677–6686 (2000).
14. Grüneis, A. *et al.* Natural orbitals for wave function based correlated calculations using a plane wave basis set. *J. Chem. Theory Comput.* **7,** 2780–2785 (2011).
15. Shiozaki, T. & Hirata, S. Communications: Explicitly correlated second-order Møller–Plesset perturbation method for extended systems. *J. Chem. Phys.* **132,** 151101 (2010).
16. Hättig, C., Klopper, W., Köhn, A. & Tew, D. P. Explicitly correlated electrons in molecules. *Chem. Rev.* **112,** 4–74 (2012).
17. Nolan, S. J., Bygrave, P. J., Allan, N. L. & Manby, F. R. Comparison of the incremental and hierarchical methods for crystalline neon. *J. Phys. Condens. Matter* **22,** 074201 (2010).
18. Kwee, H., Zhang, S. & Krakauer, H. Finite-size correction in many-body electronic structure calculations. *Phys. Rev. Lett.* **100,** 126404 (2008).
19. Manby, F. R., Stella, M., Goodpaster, J. D. & Miller, T. F. III. A simple, exact density-functional-theory embedding scheme. *J. Chem. Theory Comput.* **8,** 2564–2568 (2012).
20. Booth, G. H., Thom, A. J. W. & Alavi, A. Fermion Monte Carlo without fixed nodes: a game of life, death, and annihilation in Slater determinant space. *J. Chem. Phys.* **131,** 054106 (2009).
21. Booth, G. H. & Alavi, A. Approaching chemical accuracy using full configuration-interaction quantum Monte Carlo: a study of ionization potentials. *J. Chem. Phys.* **132,** 174104 (2010).
22. Cleland, D., Booth, G. H. & Alavi, A. Survival of the fittest: accelerating convergence in full configuration-interaction quantum Monte Carlo. *J. Chem. Phys.* **132,** 041103 (2010).
23. Cleland, D. M., Booth, G. H. & Alavi, A. A study of electron affinities using the initiator approach to full configuration interaction quantum Monte Carlo. *J. Chem. Phys.* **134,** 024112 (2011).
24. Booth, G. H., Cleland, D., Thom, A. J. W. & Alavi, A. Breaking the carbon dimer: the challenges of multiple bond dissociation with full configuration interaction quantum Monte Carlo methods. *J. Chem. Phys.* **135,** 084104 (2011).
25. Shepherd, J. J., Booth, G., Grüneis, A. & Alavi, A. Full configuration interaction perspective on the homogeneous electron gas. *Phys. Rev. B* **85,** 081103 (2012).
26. López Ríos, P. *et al.* Inhomogeneous backflow transformations in quantum Monte Carlo calculations. *Phys. Rev. E* **74,** 066701 (2006).
27. Binnie, S. J. *et al.* Bulk and surface energetics of crystalline lithium hydride: benchmarks from quantum Monte Carlo and quantum chemistry. *Phys. Rev. B* **82,** 165431 (2010).
28. Ohnishi, Y.-y. & Hirata, S. Logarithm second-order many-body perturbation method for extended systems. *J. Chem. Phys.* **133,** 034106 (2010).

29. Nemec, N., Towler, M. D. & Needs, R. J. Benchmark all-electron ab initio quantum Monte Carlo calculations for small molecules. *J. Chem. Phys.* **132,** 034111 (2010).
30. Morales, M. A., McMinis, J., Clark, B. K., Kim, J. & Scuseria, G. E. Multideterminant wave functions in quantum Monte Carlo. *J. Chem. Theor. Comput* **8,** 2181–2188 (2012).
31. Al-Saidi, W. A., Zhang, S. & Krakauer, H. Bond breaking with auxiliary-field quantum Monte Carlo. *J. Chem. Phys.* **127,** 144101 (2007).
32. Purwanto, W., Krakauer, H. & Zhang, S. Pressure-induced diamond to®-tin transition in bulk silicon: a quantum Monte Carlo study. *Phys. Rev. B* **80,** 214116 (2009).
33. Møller, C. & Plesset, M. S. Note on an approximation treatment for many-electron systems. *Phys. Rev.* **46,** 618–622 (1934).
34. Čížek, J. On the correlation problem in atomic and molecular systems. Calculation of wavefunction components in Ursell-type expansion using quantum-field theoretical methods. *J. Chem. Phys.* **45,** 4256–4266 (1966).
35. Raghavachari, K., Trucks, G. W., Pople, J. A. & Head-Gordon, M. A fifth-order perturbation comparison of electron correlation theories. *Chem. Phys. Lett.* **157,** 479–483 (1989).
36. Troyer, M. & Wiese, U.-J. Computational complexity and fundamental limitations to fermionic quantum Monte Carlo simulations. *Phys. Rev. Lett.* **94,** 170201 (2005).
37. Spencer, J. S., Blunt, N. S. & Foulkes, W. M. The sign problem and population dynamics in the full configuration interaction quantum Monte Carlo method. *J. Chem. Phys.* **136,** 054110 (2012).
38. Schimka, L., Harl, J. & Kresse, G. Improved hybrid functional for solids: the HSEsol functional. *J. Chem. Phys.* **134,** 024116 (2011).
39. Perdew, J. P., Burke, K. & Ernzerhof, M. Generalized gradient approximation made simple. *Phys. Rev. Lett.* **77,** 3865–3868 (1996).
40. Csonka, G. I. *et al.* Assessing the performance of recent density functionals for bulk solids. *Phys. Rev. B* **79,** 155107 (2009).
41. Cora, F. *et al.* in *Density Functional Theory in Inorganic Chemistry, Structure and Bonding* (eds McGrady, J. & Kaltsoyannis, N.) 171–232 (Springer, 2004).
42. Hutchings, M. & Samuelsen, E. Measurement of spin-wave dispersion in NiO by inelastic neutron-scattering and its relation to magnetic properties. *Phys. Rev. B* **6,** 3447–3461 (1972).
43. Flyvbjerg, H. & Petersen, H. G. Error estimates on averages of correlated data. *J. Chem. Phys.* **91,** 461–466 (1989).
44. Grüneis, A., Marsman, M. & Kresse, G. Second-order Møller–Plesset perturbation theory applied to extended systems. II. Structural and energetic properties. *J. Chem. Phys.* **133,** 074107 (2010).

Acknowledgements G.H.B. acknowledges support from Trinity College, Cambridge. A.G. acknowledges an APART fellowship from the Austrian Academy of Sciences. G.K. acknowledges support from the Austrian Science fund (FWF) within the SFB ViCoM (F41). A.A. acknowledges the support of the EPSRC through grants EP/I014624/1 and EP/J003867/1. The authors thank J. Spencer and A. Thom for technical contributions and discussions. Computer time on the Vienna Scientific Cluster (VSC) and HECToR (under the DEISA Extreme Computing Initiative) are acknowledged.

Author Contributions G.H.B. and A.G. contributed equally to this work. G.H.B. and A.A. developed the FCIQMC method, G.H.B. wrote the computer code, and A.G. developed the quantum-chemical methods in VASP. G.K. led the VASP project, and A.A. led the FCIQMC project.

Author Information Reprints and permissions information is available at www.nature.com/reprints. The authors declare no competing financial interests. Readers are welcome to comment on the online version of the paper. Correspondence and requests for materials should be addressed to G.H.B. (ghb24@cam.ac.uk) or A.G. (andreas.grueneis@univie.ac.at).

ARTICLE

doi:10.1038/nature11628

Autism-related deficits via dysregulated eIF4E-dependent translational control

Christos G. Gkogkas[1], Arkady Khoutorsky[1]*, Israeli Ran[2]*, Emmanouil Rampakakis[1,3], Tatiana Nevarko[1], Daniel B. Weatherill[2], Cristina Vasuta[2], Stephanie Yee[4], Morgan Truitt[5], Paul Dallaire[6], François Major[6], Paul Lasko[4], Davide Ruggero[5], Karim Nader[7], Jean-Claude Lacaille[2] & Nahum Sonenberg[1]

Hyperconnectivity of neuronal circuits due to increased synaptic protein synthesis is thought to cause autism spectrum disorders (ASDs). The mammalian target of rapamycin (mTOR) is strongly implicated in ASDs by means of upstream signalling; however, downstream regulatory mechanisms are ill-defined. Here we show that knockout of the eukaryotic translation initiation factor 4E-binding protein 2 (4E-BP2)—an eIF4E repressor downstream of mTOR—or eIF4E overexpression leads to increased translation of neuroligins, which are postsynaptic proteins that are causally linked to ASDs. Mice that have the gene encoding 4E-BP2 (*Eif4ebp2*) knocked out exhibit an increased ratio of excitatory to inhibitory synaptic inputs and autistic-like behaviours (that is, social interaction deficits, altered communication and repetitive/stereotyped behaviours). Pharmacological inhibition of eIF4E activity or normalization of neuroligin 1, but not neuroligin 2, protein levels restores the normal excitation/inhibition ratio and rectifies the social behaviour deficits. Thus, translational control by eIF4E regulates the synthesis of neuroligins, maintaining the excitation-to-inhibition balance, and its dysregulation engenders ASD-like phenotypes.

ASDs encompass unique behaviours within three domains: social interactions, communication and repetitive or restricted interests and behaviours[1]. Dysregulation of the molecular machinery controlling synaptic messenger RNA translation has been postulated to cause ASD, mainly through increased synthesis of synaptic proteins and consequently augmented connectivity[2].

A key factor required for translation initiation of the vast majority of mRNAs is the eIF4F complex that includes the cap-binding protein eIF4E, the RNA helicase eIF4A and the modular scaffolding protein eIF4G that bridges the mRNA to the ribosome[3]. 4E-BP proteins repress translation initiation by disrupting eIF4F complex formation[4]. Although most cellular mRNAs require basal amounts of eIF4E to be translated, eIF4E preferentially enhances translation of a selective group of mRNAs with extensive secondary structure at their 5′ untranslated regions (UTRs) or other sequence elements[3,5].

mTOR exists in two complexes: mTORC1, which is sensitive to the drug rapamycin, and mTORC2, which is not[3]. mTORC1 phosphorylates 4E-BPs[4], stimulating translation. Hypophosphorylated forms of 4E-BPs bind to eIF4E and inhibit translation initiation. There are three mammalian *4E-BP* paralogues encoding 4E-BP1, 4E-BP2 and 4E-BP3. 4E-BP2 is the major form in the mammalian brain and has an important role in long-lasting synaptic plasticity, learning and memory[6].

Upstream of mTOR, germline mutations in the phosphatase and tensin homologue on chromosome ten (*PTEN*) are present in about 1–5% of patients with autism[7]. *Pten* knockout mice exhibit cognitive impairment and deficits in social interactions[8], which are rescued by rapamycin[9]. Mutations in the two tuberous sclerosis complex (*TSC1* and *TSC2*) genes cause autism in a subset of patients[10]. *Tsc2*$^{+/-}$ mice display synaptic plasticity and memory deficits[11,12], which are rescued by rapamycin, and enhanced ultrasonic vocalizations (USVs) in mouse pups[13]. Deletion of *Tsc1* in cerebellar Purkinje cells in mice leads to autism-like phenotypes, which are rescued by rapamycin[14]. Recently, a *de novo* mutated eIF4GI was reported as a component of a highly interconnected network of proteins linked with sporadic autism[15]. Moreover, the mTORC1–eIF4E pathway is hyperactivated in fragile-X syndrome (FXS) patients diagnosed with autism[16]. The hypothesis that eIF4E dysregulation is causally linked to ASD is consistent with recent findings that CYFIP1 (cytoplasmic fragile-X mental retardation protein (FMRP) interacting protein 1), an eIF4E binding protein[17], is associated with ASD[18].

Aberrant information processing owing to the increased ratio of synaptic excitation to inhibition (E/I) has been proposed to cause ASD[19,20]. Children with ASD exhibit elevations in resting-state neuronal activity, supporting an E/I imbalance as a neurobiological feature of ASD[21]. Modulations of the E/I balance using optogenetics directly influence autism-like behaviours in adult mice[22]. Several ASD mouse models display E/I imbalances due to altered glutamatergic excitation or to altered GABA (γ-aminobutyric acid)-ergic inhibition, leading to an increased[23,24] or decreased[25–28] E/I ratio. The balance of excitatory and inhibitory synapses is largely controlled by the expression of adhesion molecules (neuroligins (NLGNs), neurexins) and scaffolding proteins (postsynaptic density protein 95 (PSD95), gephyrin) in neurons[29]. Overexpression, knockout or knock-in of ASD-related neuroligin mutations reveals ASD-like phenotypes with alterations in the E/I balance (summarized in Supplementary Table 1).

A causal link between dysregulated eIF4E-dependent translation and the development of ASD is lacking. Here we show that deletion of mouse *Eif4ebp2* leads to autistic-like behaviours, including social interaction deficits, altered communication and repetitive/stereotyped behaviours. We find that translation of neuroligin mRNAs is enhanced in *Eif4ebp2* knockout and eIF4E-overexpressing mice. Furthermore, unitary excitatory and inhibitory synaptic activity is increased in

---

[1] Department of Biochemistry & Goodman Cancer Research Centre, McGill University, Montreal, Quebec H3A 1A3, Canada. [2] GRSNC and Department of Physiology, Université de Montréal, Montreal, Quebec H3C 3J7, Canada. [3] JSS Medical Research Inc., Montreal, Quebec H4S 1N8, Canada. [4] Department of Biology, McGill University, Montreal, Quebec H3G 0B1, Canada. [5] School of Medicine and Department of Urology, Helen Diller Family Comprehensive Cancer Center, University of California, San Francisco, California 94158, USA. [6] Institute for Research in Immunology and Cancer, and Department of Computer Science, Université de Montréal, Montreal, Quebec H3C 3J7, Canada. [7] Department of Psychology, McGill University, Montreal, Quebec H3A 1B1, Canada.

*These authors contributed equally to this work.

Eif4ebp2 knockout mice, with excitation predominantly being augmented. Moreover, pharmacological normalization of eIF4F activity corrects the E/I imbalance and rescues the social behaviour deficits of *Eif4ebp2* knockout mice. Knockdown of *Nlgn1*, but not *Nlgn2*, in *Eif4ebp2* knockout mice reverses the defects in social interactions and rectifies the E/I imbalance. Thus, enhanced eIF4E-dependent translation of mRNAs encoding synaptic adhesion proteins, such as neuroligins, leads to an aberrant increase in excitatory synaptic function and results in ASD-like phenotypes.

Autistic-like phenotypes in *Eif4ebp2* knockout mice

As *Eif4ebp2* deletion causes increased eIF4F complex formation and activity, we postulated that it could engender ASD-related behaviours. Social interaction deficits are a salient autistic behavioural feature in humans[1]. To test our hypothesis, we used a three-chamber social arena to assess the preference of a test mouse for a social (stranger1) over a non-social (empty wire cage) stimulus or for social novelty (stranger2). In comparison with wild-type littermates, *Eif4ebp2* knockout mice displayed impaired social approach behaviour as depicted by the significantly reduced time spent in the stranger1 compartment and the decreased interaction time with stranger1 (Fig. 1a and Supplementary Fig. 2). Given the lack of preference for the social stimulus in the first portion of the three-chamber test, preference for social novelty was also reduced in *Eif4ebp2* knockout mice (Supplementary Fig. 3). Because the reduced number of entries of the *Eif4ebp2* knockout mice into the stranger1 chamber relative to wild-type littermates (Fig. 1a) might underlie an anxiety-like phenotype, we assessed anxiety levels. No differences between wild-type and *Eif4ebp2* knockout mice were found using the elevated plus maze test following the three-chamber social approach test (Supplementary Fig. 4d). Moreover, wild-type and *Eif4ebp2* knockout mice were indistinguishable in their initial exploration of the three-chamber arena (Supplementary Fig. 4a).

To characterize the social behaviour of the *Eif4ebp2* knockout mice further, we performed reciprocal social interaction tests: (1) in their home cage and (2) in a separate cage (Fig. 1b, c). Pairs of wild-type–wild-type, knockout–wild-type or knockout–knockout mice were either recorded in their home cage, or introduced to a familiar open-field environment. The knockout–wild-type and knockout–knockout pairs interacted for a shorter period of time, as compared to the wild-type–wild-type pair (Fig. 1b, c), without any difference in the total number of contacts between the wild-type–wild-type or wild-type–knockout pairs of mice. Knockout–knockout pairs initiated significantly fewer contacts (Fig. 1b, c). This discrepancy can be attributed to the larger number of contacts initiated by the wild-type mice in the wild-type–knockout pairs (Supplementary Fig. 2h). Anxiety is not a confounding factor for these behaviours, as demonstrated by the similar anxiety levels in wild-type and *Eif4ebp2* knockout mice in the elevated plus maze test (Supplementary Fig. 4b, c). To assess stereotyped/repetitive behaviours, we recorded self-grooming and used a marble-burying test. *Eif4ebp2* knockout mice displayed longer bouts and total duration of self-grooming (Fig. 1d and Supplementary Fig. 2i) and buried significantly more marbles than wild-type littermates (Fig. 1e).

Isolation-induced USVs emitted from mouse pups separated from their mother are a quantifiable behaviour considered relevant to ASD[30]. $Tsc2^{+/-}$ pups display enhanced USVs (increased number of calls, duration, peak amplitude and peak frequency)[13,14]. We recorded USVs in *Eif4ebp2* knockout and wild-type littermate pups after separation from their dams (Fig. 1f–h). *Eif4ebp2* knockout pups emitted a higher number of calls (Fig. 1f) with a longer duration (Fig. 1g) at all postnatal days relative to wild-type littermates. Furthermore, the

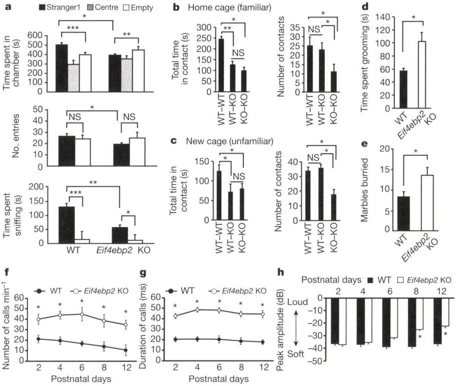


Figure 1 | Social interaction deficits, repetitive behaviour and elevated USVs in *Eif4ebp2* knockout mice. a, Three-chamber social interaction test showing time spent in each chamber, number of entries and time spent sniffing the wire cage; $n = 12$ for each group; $***P < 0.001$, $**P < 0.01$, $*P < 0.03$, two-way analysis of variance (ANOVA) with Bonferroni's post-hoc test. KO, knockout; NS, not significant; WT, wild type. **b**, **c**, Home cage (**b**) and reciprocal social interactions (**c**) tests showing total time in contact and number of contacts; $n = 18$ per group; $**P < 0.02$, $*P < 0.03$; one way ANOVA with Bonferroni's post-hoc test. **d**, Self-grooming test showing total time spent grooming. **e**, Marble-burying test showing marbles buried; $n = 12$ per group; $*P < 0.03$; Student's *t*-test. **f–h**, Isolation-induced USVs at various postnatal days. Shown are number of calls per min (**f**), duration of calls (**g**) and amplitude of USVs (**h**) in pups; $n = 12$ per group; $*P < 0.05$; two-way ANOVA, with Bonferroni's post-hoc test. All error bars show mean ± s.e.m.

peak amplitude of the USVs was higher on postnatal days 8 and 12 (Fig. 1h). Taken together, our data demonstrate that genetic deletion of *Eif4ebp2*, mimicking eIF4E upregulation, engenders ASD-like behaviours.

Increased eIF4E-dependent translation of neuroligins

We hypothesized that the ASD-like phenotypes of the *Eif4ebp2* knockout mice are a result of altered translation of a subset of mRNAs, the initiation of which is stimulated by eIF4E activity. To test this hypothesis, we performed polysome profiling, which measures translation initiation rates of hippocampal lysates from *Eif4ebp2* knockout and eIF4E-overexpressing (βT-*Eif4e*)[31] mice. The polysome profile was not significantly altered in *Eif4ebp2* knockout mice as compared to wild-type littermates (Fig. 2a and Supplementary Fig. 5d). This is in agreement with the lack of changes in [^{35}S]Met/Cys incorporation (Supplementary Fig. 5a, b) and previous reports[32]. We then examined the polysome distribution of 24 mRNAs coding for proteins known to be associated with ASD (Supplementary Table 2)[33,34]. These include mRNAs encoding the family of neuroligins and neurexins, SHANK2 and SHANK3 (SH3 and multiple ankyrin repeat domains protein) and MeCP2 (methyl-CpG-binding protein). Notably, out of the 24 mRNAs examined, only mRNA for *Nlgn1*, *Nlgn2*, *Nlgn3* and *Nlgn4* was shifted towards heavier polysome fractions in the *Eif4ebp2* knockout mice relative to wild-type littermates (Fig. 2c), indicating increased translation initiation. Neither mRNA levels of *Nlgn1*–*Nlgn4* nor those of any other mRNA were different between *Eif4ebp2* knockout and wild-type mice (Supplementary Fig. 6c), thus excluding transcriptional effects. As expected, protein amounts of all four neuroligins were increased in crude and synaptosomal extracts from *Eif4ebp2* knockout hippocampi (Fig. 2e, g). Elevated translation, but not transcription, of neuroligin mRNAs was also observed in polysome profiles of βT-*Eif4e* mice (Fig. 2d and Supplementary Fig. 6c) with no differences in translation or transcription of other mRNAs (Supplementary Figs 6c and 7c). Accordingly, protein levels of all four neuroligins were increased in the synaptosomal fractions from βT-*Eif4e* mice relative to wild type (Fig. 2f, h). The increased translation of neuroligins, observed only in synaptosomal fractions of βT-*Eif4e* mice, can be explained by the mild overexpression (by 29%) of eIF4E (Fig. 2f). In contrast, no changes were observed in the translation of mRNAs encoding neurexins (Supplementary Fig. 6) or PSD95 (*Dlg4* gene; Fig. 2c–f) in *Eif4ebp2* knockout or βT-*Eif4e* versus wild-type mice. Translation of mRNAs coding for gephyrin (*Gphn*), SHANK2, SHANK3 and SAPAP3 (PSD95/synapse associated protein 90-binding protein 3) (*Dlgap3*) proteins, which orchestrate the formation of the postsynaptic scaffolding[29], was not changed in *Eif4ebp2* knockout or in the βT-*Eif4e* mice (Supplementary Fig. 7). In conclusion, relief of translational suppression by loss of 4E-BP2 or by overexpression of eIF4E selectively enhances the synthesis of neuroligins, causing an imbalance between these adhesion proteins relative to scaffolding proteins.

Increased synaptic activity in *Eif4ebp2* knockouts

An altered E/I ratio was reported in several ASD mouse models[23–28]. Therefore, we examined the effects of *Eif4ebp2* knockout on unitary excitatory and inhibitory synaptic transmission in acute hippocampal

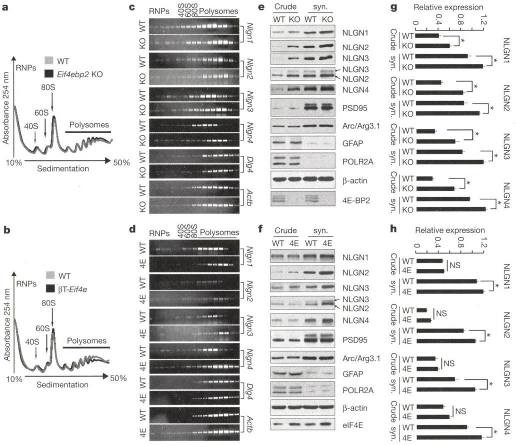


Figure 2 | **Enhanced eIF4E-dependent translation of neuroligin mRNAs.** **a**, **b**, Polysome profiles from hippocampal lysates of wild-type and *Eif4ebp2* knockout mice (**a**) and wild-type and βT-*Eif4e* mice (**b**). Positions of 40S, 60S and 80S ribosome peaks and polysomes are indicated. **c**, **d**, RT–PCR on RNA extracted from polysome fractions from wild-type and *Eif4ebp2* knockout (**c**) and wild-type and βT-*Eif4e* (**d**; 4E) hippocampi. Representative gel images are shown; $n = 4$ per group. **e**, **f**, Representative immunoblots of crude and synaptosomal (syn.) fractions from hippocampal lysates of wild-type or *Eif4ebp2* knockout (**e**) and wild-type or βT-*Eif4e* (**f**; 4E) mice, probed with antibodies against the indicated proteins. **g**, **h**, Quantification of immunoblots from **e** and **f** for neuroligins; $n = 4$ per group; $*P < 0.05$; Student's *t*-test. All error bars show mean ± s.e.m.

slices. AMPA (α-amino-3-hydroxy-5-methyl-4-isoxazole propionic acid) receptor-mediated miniature excitatory postsynaptic currents (mEPSCs) were increased both in amplitude and frequency in pyramidal neurons of *Eif4ebp2* knockout relative to wild-type mice (Fig. 3a, b), indicating that 4E-BP2 de-repression augments transmission at excitatory synapses in acute hippocampal slices. We next examined whether transmission was similarly altered at inhibitory synapses. GABA$_A$ receptor-mediated miniature inhibitory postsynaptic currents (mIPSCs) were significantly greater in amplitude, but not in frequency (unlike mEPSCs), in pyramidal cells of *Eif4ebp2* knockout relative to wild-type slices (Fig. 3c, d). To study the differential effect on excitatory versus inhibitory transmission, we measured the total synaptic charge transfer for mEPSCs and mIPSCs over a 10-min period—a parameter reflecting both amplitude and frequency of miniature synaptic events. The total charge transfer was significantly increased for both mEPSCs and mIPSCs in *Eif4ebp2* knockout pyramidal cells (Fig. 3e). However, the normalized increase of total charge transfer in *Eif4ebp2* knockout relative to wild-type mice was significantly larger for mEPSCs than for mIPSCs (Fig. 3f). We also recorded minimal stimulation-evoked EPSCs and IPSCs from CA1 pyramidal neurons in wild-type and *Eif4ebp2* knockout slices and observed an increase in both EPSC and IPSC amplitude, but no change in paired-pulse ratio (Supplementary Fig. 8). Moreover, we observed a greater increase in vesicular glutamate transporter (VGLUT) than vesicular GABA transporter (VGAT) puncta and protein amounts in the CA1 dendritic layer of *Eif4ebp2* knockout hippocampal slices, relative to wild-type littermates (Supplementary Fig. 9a–g). Finally, spine density is increased in CA1 pyramidal cell dendrites in *Eif4ebp2* knockout mice (Supplementary Fig. 9h, i). Thus, 4E-BP2 de-repression results in altered E/I balance.

Pharmacological rescue of ASD-like phenotypes

To bolster the conclusion that translation initiation through eIF4E is critical in altering synaptic function and thus causing ASD, we used the selective inhibitor 4EGI-1 (ref. 35), which prevents eIF4E binding to eIF4G, to test whether it would rescue autistic-like behaviours in *Eif4ebp2* knockout mice. First, we performed cap-column pull-down assays with hippocampal lysates to assess the status of the eIF4F complex. As expected, increased amounts of eIF4G and eIF4E were bound to the m^{7}GDP-cap analogue matrix in *Eif4ebp2* knockout mice, as compared to wild-type mice (Fig. 4b). Infusion of 4EGI-1 reduced the amounts of cap-bound eIF4G and eIF4E to wild-type levels (Fig. 4b). Notably, neuroligin protein levels were also reduced to wild-type levels after 4EGI-1 treatment (Fig. 4b).

4EGI-1 blocks long-term memory consolidation but not reconsolidation in mice[36]. In *Eif4ebp2* knockout mice a single tetanic train elicits late-phase long-term potentiation (L-LTP)[6], indicating a lower threshold for L-LTP induction compared to wild-type mice. Bath application of 4EGI-1 in acute hippocampal slices inhibited the facilitation of L-LTP induction in *Eif4ebp2* knockout mice (Supplementary Fig. 10c). Moreover, application of 4EGI-1 (Fig. 4a) to hippocampal slice cultures from *Eif4ebp2* knockout mice reversed the increase in mEPSC amplitude and frequency (Fig. 4c), reversed the increase in mIPSC amplitude (Fig. 4d), decreased the total mEPSC and mIPSC charge transfer (Fig. 4e), and restored the E/I balance.

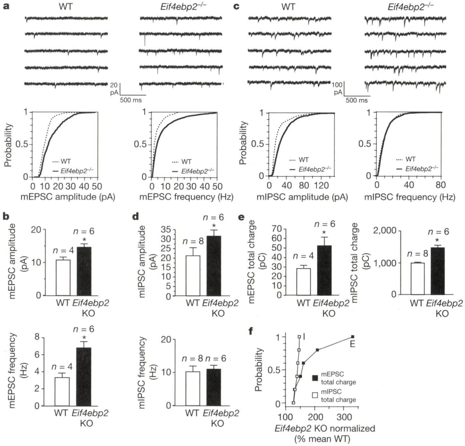


Figure 3 | Excitation is increased to a larger extent than inhibition in *Eif4ebp2* knockout mice. **a**, **c**, Representative traces (top) and cumulative distribution plots (bottom) of mEPSCs (**a**) and mIPSCs (**c**) from CA1 pyramidal neurons in acute slices from wild-type and *Eif4ebp2*$^{-/-}$ mice, illustrating changes in amplitude and frequency. **b**, **d**, Bar graphs showing mEPSC (**b**) and mIPSC (**d**) amplitude and frequency (n = number of cells). **e**, Bar graphs showing the synaptic total charge transfer in *Eif4ebp2* knockout relative to wild-type slices. For **a**–**e**, $*P < 0.05$; Student's t-test. **f**, Relative changes in mEPSC and mIPSC total charge transfer, normalized to the mean wild-type value, for each neuron from *Eif4ebp2* knockout slices; Kolmogorov–Smirnov test, $P = 0.005$, KS statistic = 0.916; E, excitatory; I, inhibitory. All error bars show mean ± s.e.m.

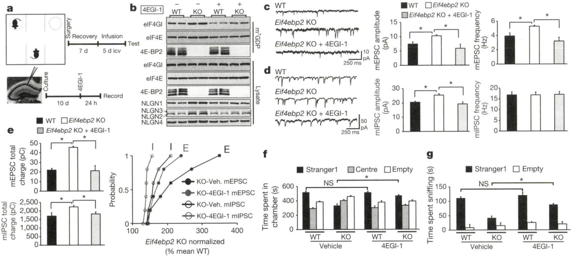


Figure 4 | Rescue of excitatory/inhibitory synaptic activity imbalance and social deficits in *Eif4ebp2* knockout mice by inhibiting the eIF4E–eIF4G interaction. **a**, Schematic of 4EGI-1 treatment in the three-chamber test (top) and slice-culture recordings (bottom). icv, intracerebroventricular infusion. **b**, Immunoblot analysis of cap column pull down in *Eif4ebp2* knockout and wild-type mice treated with 4EGI-1 or vehicle. Total and m^7GDP bound extracts were probed for eIF4GI, eIF4E and 4E-BP2; $n = 4$ (quantification in Supplementary Fig. 16a). **c**, **d**, Effects of 4EGI-1 on mEPSC (**c**) or mIPSC (**d**) amplitude and frequency in *Eif4ebp2* knockout slices. **e**, Effects of 4EGI-1 on total charge transfer of mEPSCs (top) and mIPSCs (bottom) in *Eif4ebp2* knockout slices. Right: relative changes in mEPSC and mIPSC total charge transfer, normalized to the mean wild-type value, for *Eif4ebp2* knockout neurons. Kolmogorov–Smirnov test. E, excitatory; I, inhibitory. For **c**–**e**, $n = 6$ per group; $*P < 0.05$; one-way ANOVA with Bonferroni's post-hoc test. **f**, **g**, Rescue effects of 4EGI-1 infusion in the three-chamber test showing time spent in chambers (**f**) and time spent sniffing the wire cage (**g**) in *Eif4ebp2* knockout and wild-type mice. For **f**, **g**, $n = 12$ per group; $*P < 0.03$; two-way ANOVA, with Bonferroni's post-hoc test. All error bars show mean ± s.e.m.

Next, we examined whether 4EGI-1 could rescue the social behaviour deficits in *Eif4ebp2* knockout mice. We used intracerebroventricular cannulation to infuse 4EGI-1 daily for 5 days (Fig. 4a). Mice were then subjected to the three-chamber social interaction test. Whereas a sub-threshold concentration of 4EGI-1 had no effect on wild-type mice (Fig. 4f and Supplementary Fig. 10b), it reversed social interaction deficits of *Eif4ebp2* knockout mice, inasmuch as 4EGI-1-treated *Eif4ebp2* knockout mice spent significantly more time in the stranger1 chamber (Fig. 4f), whereas the number of entries for the *Eif4ebp2* knockout mice infused with 4EGI-1 was restored to wild-type levels (Supplementary Fig. 12a). Also, the interaction time with stranger1 was increased compared to vehicle-treated knockout mice (Fig. 4g). These results establish causality between inactivation of the eIF4F complex and rescue of the social interaction impairments in the *Eif4ebp2* knockout mice.

Nlgn1 knockdown rescues ASD-like *Eif4ebp2* phenotypes

NLGN1, which is present exclusively at excitatory synapses, promotes excitatory synaptic transmission[37,38]. In *Eif4ebp2* knockout mice *Nlgn1* mRNA translation is enhanced concomitant with increased E/I ratio (Figs 2 and 3). We therefore predicted that reduction of NLGN1 levels in *Eif4ebp2* knockout mice would restore the E/I balance and reverse the autism-related deficits. To investigate this, we used short interfering RNAs (siRNAs) against *Nlgn1* in hippocampal pyramidal cells in slice cultures (Fig. 5a). Knockdown of *Nlgn1* rescued the amplitude and frequency of mEPSCs in *Eif4ebp2* knockout pyramidal cells (Fig. 5c), without affecting mIPSCs (Fig. 5d). As expected, mEPSC total charge transfer in pyramidal cells of *Eif4ebp2* knockout mice transfected with *Nlgn1* siRNA was restored to wild-type levels (Fig. 5e), whereas mIPSC total charge transfer was unaffected (Fig. 5e). Next, we knocked down *Nlgn2*, a regulator of inhibitory synapses. In contrast to NLGN1, knockdown of *Nlgn2* rescued the increase in mIPSC amplitude in *Eif4ebp2* knockout pyramidal cells (Fig. 5d), without affecting the amplitude or frequency of mEPSCs (Fig. 5c). The increase in total charge transfer of mIPSCs was also attenuated in *Nlgn2*-siRNA-transfected *Eif4ebp2* knockout pyramidal cells, whereas mEPSC total charge transfer remained unaltered (Fig. 5e).

We next subjected wild-type or *Eif4ebp2* knockout mice to the three-chamber social interaction test after intracerebroventricular injection of shRNAs (Fig. 5a and Supplementary Fig. 11). Notably, NLGN1 or NLGN2 protein levels decreased to wild-type levels in *Eif4ebp2* knockout mice injected with *Nlgn1* or *Nlgn2* shRNA, respectively, whereas protein levels of the other neuroligins were not changed (Fig. 5b). We selected a concentration of lentiviral shRNA (Supplementary Fig. 11) that had no effect in wild-type mice (Fig. 5f, g). Importantly, this sub-threshold concentration of *Nlgn1* shRNA partially rescued the social interaction impairment in *Eif4ebp2* knockout mice as follows: *Eif4ebp2* knockout mice injected with *Nlgn1* shRNA spent significantly more time in the stranger1 chamber (Fig. 5f), entered the chamber more times (Supplementary Fig. 12b), and interacted more with stranger1 (Fig. 5g) relative to control knockout mice. In contrast, knockdown of *Nlgn2* exacerbated the social deficits in *Eif4ebp2* knockout mice, as shown by the reduced time spent in the stranger1 chamber (Fig. 5f), and the decreased amount of time spent interacting with stranger1 (Fig. 5g) as compared to control knockout mice. We recapitulated these findings with additional shRNAs against *Nlgn1* or *Nlgn2* (Supplementary Fig. 13). In summary, NLGN1 knockdown reverses changes at excitatory synapses and partially rescues the social interaction deficits in *Eif4ebp2* knockout mice, whereas NLGN2 knockdown weakens inhibitory synapses and exacerbates the phenotype, thus establishing a strong link between eIF4E-dependent translational control of neuroligins, E/I balance and the development of ASD-like phenotypes.

Discussion

We show that 4E-BP2 deficiency in mice causes autistic-like phenotypes that markedly resemble behavioural hallmarks of ASD. *Eif4ebp2* knockout mice display a previously undocumented deficit in social interactions and exhibit repetitive/stereotyped behaviours. Communication between *Eif4ebp2* knockout pups and their dams is altered, as demonstrated by enhanced ultrasonic vocalizations. This phenotype is similar to that reported for $Tsc2^{+/-}$ (ref. 13) or conditional *Tsc2* knockout[14] mice, which could conceivably be explained by increased eIF4E activity. These newly discovered autistic-like phenotypes of *Eif4ebp2* knockout mice are neither confounded by anxiety

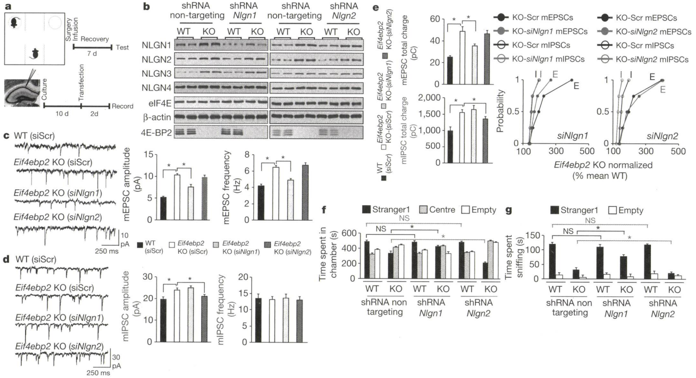


Figure 5 | Knockdown of neuroligin 1 rescues the excitatory/inhibitory synaptic activity imbalance and social deficits in *Eif4ebp2* knockout mice. **a**, Schematic of the three-chamber test in mice intracerebroventricularly injected with lentiviruses expressing shRNAs (top) and recordings from siRNA-transfected slice cultures (bottom). **b**, Western blot analysis of hippocampal lysates from wild-type or *Eif4ebp2* knockout mice injected with lentiviruses expressing shRNAs against a non-targeting sequence or *Nlgn1* or *Nlgn2*. Extracts were probed for NLGN1– NLGN4, β-actin, eIF4E and 4E-BP2; $n = 4$ (quantification in Supplementary Fig. 16b, c). **c**, **d**, Effects of *Nlgn1* or *Nlgn2* knockdown on mEPSC (**c**) or mIPSC (**d**) amplitude and frequency in transfected pyramidal cells. *siNlgn1* or *siNlgn2*, siRNA against *Nlgn1* or *Nlgn2*; siScr, scrambled siRNA. **e**, mEPSC (top) or mIPSC (bottom) total charge transfer in *Nlgn1* or *Nlgn2* knockdown cells. Right: relative changes in mEPSC and mIPSC total charge transfer, normalized to the mean wild-type value, for each neuron from *Eif4ebp2* knockout slices; Kolmogorov–Smirnov test. For **c**–**e**, $n = 5$ per group; $*P < 0.05$; one-way ANOVA with Bonferroni's post-hoc test. **f**, **g**, Knockdown of *Nlgn1* or *Nlgn2* in the three-chamber test showing time spent in chambers (**f**) and time spent sniffing the wire cage (**g**). For **f**, **g**, $n = 12$ per group; $*P < 0.02$; two-way ANOVA, with Bonferroni's post-hoc test. All error bars show mean ± s.e.m.

nor by the previously reported long-term memory deficits[6,39], as short-term memory circuits are intact. Thus, *Eif4ebp2* knockout mice constitute a novel model for mTORC1/eIF4E-dependent autism-like phenotypes due to dysregulated translational control. In accordance with this model, a point mutation in the promoter region of the *EIF4E* gene is associated with increased activity of eIF4E in autistic siblings from two unrelated families[40].

In two models of enhanced eIF4E activity (*Eif4ebp2* knockout and eIF4E-overexpression mice), we show that neuroligin mRNAs are translated more efficiently. These results highlight the important role of 4E-BP2 and eIF4E in postsynaptic translational control of neuronal mRNAs. Increased mTOR signalling to its downstream effectors promotes the preferential translation of a subset of mRNAs referred to as 'eIF4E sensitive', harbouring extensive secondary structure elements in their 5′ UTRs[5,41]. Reporter mRNAs containing the full-length 5′ UTRs of *Nlgn1* and *Nlgn2*, but not *Dlg4* or *Actb*, fused to luciferase, are better translated in cells with elevated mTORC1 signalling (*Pten*$^{+/-}$, *Tsc2* knockout) and/or enhanced cap-dependent translation (*Eif4ebp2* knockout, eIF4E overexpression) as compared to parental cells (Supplementary Fig. 14). Interestingly, the 5′ UTRs of neuroligin mRNAs possess a repeated structural element that is absent from the other mRNA 5′ UTRs (Supplementary Fig. 15). These secondary structure repeats might account for the preferential increase in translation of neuroligins in response to increased eIF4F activity.

Mutations and copy number variation in neuroligins have been associated with ASD[42]. In *Eif4ebp2* knockout mice, increased amounts of NLGN1– NLGN4 lead to a disproportionate ratio of neuroligins to PSD95 (in excitatory synapses) or to gephyrin (in inhibitory synapses) because the general composition of the postsynaptic scaffold (other than neuroligins) is not altered by *Eif4ebp2* deletion or similarly overexpression of eIF4E. An altered E/I synaptic ratio is a prominent feature in ASD and in mouse models of autism[19,21]. The relative impact of *Eif4ebp2* on excitation is greater than on inhibition, thus engendering a shift in the E/I balance. In *Eif4ebp2* knockout mice, pharmacological inhibition of the eIF4E–eIF4G interaction using the selective inhibitor 4EGI-1 reduces eIF4F activity to wild-type levels, restores the L-LTP threshold and rescues the E/I imbalance phenotype. In *Eif4ebp2* knockout mice, knockdown of *Nlgn1* restores the E/I balance and improves social approach behaviour to wild-type levels. In contrast, *Nlgn2* knockdown increases the E/I imbalance and exacerbates the social behaviour deficits. These results establish a link between eIF4E translational control of neuroligins, E/I balance and ASD-like phenotypes (Supplementary Fig. 1).

This novel regulatory mechanism should provide an impetus to investigate the role of eIF4E-mediated translation in ASD further and to explore therapeutic avenues for the pharmacological modulation of eIF4F activity (for example, with 4EGI-1 and its derivatives) or the function of neuroligins.

METHODS SUMMARY

All procedures are in accordance with the collaborating institutions' Animal Care Committee guidelines. Experimenters were blinded to the genotype during testing and scoring. All data are presented as mean ± s.e.m. (error bars). Statistical results, along with tests used, are summarized in Supplementary Table 4. Statistical significance was set a priori at 0.05.

Full Methods and any associated references are available in the online version of the paper.

Received 10 January; accepted 28 September 2012.
Published online 21 November 2012.

1. Fombonne, E. Epidemiology of pervasive developmental disorders. *Pediatr. Res.* **65,** 591–598 (2009).
2. Kelleher, R. J. & Bear, M. F. The autistic neuron: troubled translation? *Cell* **135,** 401–406 (2008).
3. Hay, N. & Sonenberg, N. Upstream and downstream of mTOR. *Genes Dev.* **18,** 1926–1945 (2004).
4. Pause, A. *et al.* Insulin-dependent stimulation of protein synthesis by phosphorylation of a regulator of 5′-cap function. *Nature* **371,** 762–767 (1994).
5. Koromilas, A. E., Lazaris-Karatzas, A. & Sonenberg, N. mRNAs containing extensive secondary structure in their 5′ non-coding region translate efficiently in cells overexpressing initiation factor eIF-4E. *EMBO J.* **11,** 4153–4158 (1992).
6. Banko, J. L. *et al.* The translation repressor 4E-BP2 is critical for eIF4F complex formation, synaptic plasticity, and memory in the hippocampus. *J. Neurosci.* **25,** 9581–9590 (2005).
7. Zhou, J. & Parada, L. F. PTEN signaling in autism spectrum disorders. *Curr. Opin. Neurobiol.* http://dx.doi.org/10.1016/j.conb.2012.05.004 (2012).
8. Kwon, C. H. *et al.* Pten regulates neuronal arborization and social interaction in mice. *Neuron* **50,** 377–388 (2006).
9. Zhou, J. *et al.* Pharmacological inhibition of mTORC1 suppresses anatomical, cellular, and behavioral abnormalities in neural-specific Pten knock-out mice. *J. Neurosci.* **29,** 1773–1783 (2009).
10. Jeste, S. S., Sahin, M., Bolton, P., Ploubidis, G. B. & Humphrey, A. Characterization of autism in young children with tuberous sclerosis complex. *J. Child Neurol.* **23,** 520–525 (2008).
11. Auerbach, B. D., Osterweil, E. K. & Bear, M. F. Mutations causing syndromic autism define an axis of synaptic pathophysiology. *Nature* **480,** 63–68 (2011).
12. Ehninger, D. *et al.* Reversal of learning deficits in a *Tsc2*$^{+/-}$ mouse model of tuberous sclerosis. *Nature Med.* **14,** 843–848 (2008).
13. Young, D. M., Schenk, A. K., Yang, S. B., Jan, Y. N. & Jan, L. Y. Altered ultrasonic vocalizations in a tuberous sclerosis mouse model of autism. *Proc. Natl Acad. Sci. USA* **107,** 11074–11079 (2010).
14. Tsai, P. T. *et al.* Autistic-like behaviour and cerebellar dysfunction in Purkinje cell *Tsc1* mutant mice. *Nature* **488,** 647–651 (2012).
15. O'Roak, B. J. *et al.* Sporadic autism exomes reveal a highly interconnected protein network of *de novo* mutations. *Nature* **485,** 246–250 (2012).
16. Hoeffer, C. A. *et al.* Altered mTOR signaling and enhanced CYFIP2 expression levels in subjects with fragile X syndrome. *Genes Brain Behav.* **11,** 332–341 (2012).
17. Napoli, I. *et al.* The fragile X syndrome protein represses activity-dependent translation through CYFIP1, a new 4E-BP. *Cell* **134,** 1042–1054 (2008).
18. Nowicki, S. T. *et al.* The Prader-Willi phenotype of fragile X syndrome. *J. Dev. Behav. Pediatr.* **28,** 133–138 (2007).
19. Rubenstein, J. L. & Merzenich, M. M. Model of autism: increased ratio of excitation/inhibition in key neural systems. *Genes Brain Behav.* **2,** 255–267 (2003).
20. Uhlhaas, P. J. & Singer, W. Neuronal dynamics and neuropsychiatric disorders: toward a translational paradigm for dysfunctional large-scale networks. *Neuron* **75,** 963–980 (2012).
21. Cornew, L., Roberts, T. P., Blaskey, L. & Edgar, J. C. Resting-state oscillatory activity in autism spectrum disorders. *J. Autism Dev. Disord.* **42,** 1884–1894 (2012).
22. Yizhar, O. *et al.* Neocortical excitation/inhibition balance in information processing and social dysfunction. *Nature* **477,** 171–178 (2011).
23. Luikart, B. W. *et al.* Pten knockdown *in vivo* increases excitatory drive onto dentate granule cells. *J. Neurosci.* **31,** 4345–4354 (2011).
24. Bateup, H. S., Takasaki, K. T., Saulnier, J. L., Denefrio, C. L. & Sabatini, B. L. Loss of Tsc1 *in vivo* impairs hippocampal mGluR-LTD and increases excitatory synaptic function. *J. Neurosci.* **31,** 8862–8869 (2011).
25. Peça, J. *et al.* Shank3 mutant mice display autistic-like behaviours and striatal dysfunction. *Nature* **472,** 437–442 (2011).
26. Schmeisser, M. J. *et al.* Autistic-like behaviours and hyperactivity in mice lacking ProSAP1/Shank2. *Nature* **486,** 256–260 (2012).
27. Won, H. *et al.* Autistic-like social behaviour in Shank2-mutant mice improved by restoring NMDA receptor function. *Nature* **486,** 261–265 (2012).
28. Chao, H. T. *et al.* Dysfunction in GABA signalling mediates autism-like stereotypies and Rett syndrome phenotypes. *Nature* **468,** 263–269 (2010).
29. Levinson, J. N. & El-Husseini, A. Building excitatory and inhibitory synapses: balancing neuroligin partnerships. *Neuron* **48,** 171–174 (2005).
30. Scattoni, M. L., Crawley, J. & Ricceri, L. Ultrasonic vocalizations: a tool for behavioural phenotyping of mouse models of neurodevelopmental disorders. *Neurosci. Biobehav. Rev.* **33,** 508–515 (2009).
31. Ruggero, D. *et al.* The translation factor eIF-4E promotes tumor formation and cooperates with c-Myc in lymphomagenesis. *Nature Med.* **10,** 484–486 (2004).
32. Graff, J. R. *et al.* Therapeutic suppression of translation initiation factor eIF4E expression reduces tumor growth without toxicity. *J. Clin. Invest.* **117,** 2638–2648 (2007).
33. Glessner, J. T. *et al.* Autism genome-wide copy number variation reveals ubiquitin and neuronal genes. *Nature* **459,** 569–573 (2009).
34. Wang, K. *et al.* Common genetic variants on 5p14.1 associate with autism spectrum disorders. *Nature* **459,** 528–533 (2009).
35. Moerke, N. J. *et al.* Small-molecule inhibition of the interaction between the translation initiation factors eIF4E and eIF4G. *Cell* **128,** 257–267 (2007).
36. Hoeffer, C. A. *et al.* Inhibition of the interactions between eukaryotic initiation factors 4E and 4G impairs long-term associative memory consolidation but not reconsolidation. *Proc. Natl Acad. Sci. USA* **108,** 3383–3388 (2011).
37. Chubykin, A. A. *et al.* Activity-dependent validation of excitatory versus inhibitory synapses by neuroligin-1 versus neuroligin-2. *Neuron* **54,** 919–931 (2007).
38. Dahlhaus, R. *et al.* Overexpression of the cell adhesion protein neuroligin-1 induces learning deficits and impairs synaptic plasticity by altering the ratio of excitation to inhibition in the hippocampus. *Hippocampus* **20,** 305–322 (2010).
39. Banko, J. L. *et al.* Behavioral alterations in mice lacking the translation repressor 4E-BP2. *Neurobiol. Learn. Mem.* **87,** 248–256 (2007).
40. Neves-Pereira, M. *et al.* Deregulation of EIF4E: a novel mechanism for autism. *J. Med. Genet.* **46,** 759–765 (2009).
41. Larsson, O. *et al.* Distinct perturbation of the translatome by the antidiabetic drug metformin. *Proc. Natl Acad. Sci. USA* **109,** 8977–8982 (2012).
42. Levy, D. *et al.* Rare *de novo* and transmitted copy-number variation in autistic spectrum disorders. *Neuron* **70,** 886–897 (2011).

Supplementary Information is available in the online version of the paper.

Acknowledgements This work was supported by the Canadian Institutes of Health Research (N.S., MOP-114994; J.-C.L., MOP-10848; P.D. and F.M., MOP-93679; and P.L. and N.S., MOP-44050), Autism Speaks (Grant 7109 to N.S.), and the Fonds de la Recherche en Santé du Québec (J.-C.L. FRSQ; Groupe de Recherche sur le Système Nerveux Central), and the National Institutes of Health (D.R.; NIH R01 CA154916 and NIH R01 CA140456). D.R. is a Leukemia & Lymphoma Society Scholar. J.-C.L. is the recipient of the Canada Research Chair in Cellular and Molecular Neurophysiology. I.R. was supported by a Fellowship of the Savoy Foundation. We thank Y. Svitkin, A. Parsyan, E. Petroulakis, R. Karni and V. Polunovski for advice; K. Gamache, A. Sylvestre, S. Perreault, C. Lister and I. Harvey for technical assistance; T. Alain for assistance with lentiviral titration; S. Hamdani for assistance with USVs; and W. Sossin and P. Skehel for critical reading of the manuscript.

Author Contributions C.G.G. and N.S. conceived and designed this study, wrote the manuscript, supervised and coordinated the project. C.G.G. carried out behavioural, biochemical and imaging experiments, data and statistical analysis; C.G.G., A.K., I.R., D.B.W. and C.V. carried out electrophysiology experiments and data analysis; T.N. and S.Y. conducted biochemical experiments and data analysis; E.R. and I.R. carried out statistical analysis; P.D. and F.M. carried out bioinformatics analysis; M.T. and D.R. provided critical insight and reagents, and edited the manuscript; P.L. supervised the project and edited the manuscript; and K.N. contributed to the design of behavioural experiments, edited the manuscript and supervised the project; J-C.L. supervised, conceived and designed the electrophysiological experiments, edited the manuscript and supervised the project.

Author Information Reprints and permissions information is available at www.nature.com/reprints. The authors declare no competing financial interests. Readers are welcome to comment on the online version of the paper. Correspondence and requests for materials should be addressed to N.S. (nahum.sonenberg@mcgill.ca) or J.-C.L. (jean-claude.lacaille@umontreal.ca).

doi:10.1038/nature11746

Pulsed accretion in a variable protostar

James Muzerolle[1], Elise Furlan[2], Kevin Flaherty[3], Zoltan Balog[4] & Robert Gutermuth[5]

Periodic increases in luminosity arising from variable accretion rates have been predicted for some pre-main-sequence close binary stars as they grow from circumbinary disks[1–3]. The phenomenon is known as pulsed accretion and can affect the orbital evolution and mass distribution of young binaries[2,4], as well as the potential for planet formation[5,6]. Accretion variability is a common feature of young stars, with a large range of amplitudes and timescales as measured from multi-epoch observations at optical[7,8] and infrared[9–13] wavelengths. Periodic variations consistent with pulsed accretion have been seen in only a few young binaries via optical accretion tracers[14–16], albeit intermittently with accretion luminosity variations ranging from zero to 50 per cent from orbit to orbit. Here we report that the infrared luminosity of a young protostar (of age about 10^5 years) increases by a factor of ten in roughly one week every 25.34 days. We attribute this to pulsed accretion associated with an unseen binary companion. The strength and regularity of this accretion signal is surprising; it may be related to the very young age of the system, which is a factor of ten younger than the other pulsed accretors previously studied.

We obtained multi-epoch mid-infrared (MIR) observations of the star forming region IC 348 using the Spitzer Space Telescope. Among the roughly 300 pre-main-sequence objects in the cluster, the protostar LRLL 54361 (which we refer to here as L54361) exhibits by far the largest MIR flux variability. We have a total of 81 separate observations of L54361 taken with all three instruments on board Spitzer. The multi-epoch spectral energy distribution (SED) is shown in Fig. 1. The measured bolometric luminosity of the system ranges from about 0.2 to 2.7 solar luminosities ($L_\odot$). The spectral shape remains relatively constant over this range, apart from slightly bluer MIR colours at higher luminosities.

The photometric light curve indicates that the flux variations occur repeatedly throughout the seven-year span of our observations. The two longest contiguous sets of photometry (Fig. 2) reveal a strong pulse signature in which the flux increases by about two magnitudes in as little as a few days, followed by a longer exponential decay over the following few weeks. The combined photometric data set suggests that the variability of L54361 is periodic in nature; the pulse shape revealed by the contiguous warm Spitzer photometry appears in the older data and at other wavelengths, albeit with variations in the pulse width and peak flux. Using several statistical tests, we find that the flux peaks repeat with a robust period of 25.34 ± 0.01 d.

Follow-up multi-epoch imaging taken with the Hubble Space Telescope at near-infrared wavelengths reveals spatially resolved scattered light structures associated with L54361 (Fig. 3). The central source varies with almost the same amplitude and light curve shape at 1.6 μm as seen in the Spitzer data, and the peak occurs exactly as expected given the previously determined period. The geometry of the scattered light is similar to that of other protostars[17], and is probably produced by cavities carved out of an infalling envelope by one or more outflows. The apparent motion of the scattered light indicates a light echo produced as the pulse peak light travels through the outflow cavities, and suggests that the source of the illumination is relatively isotropic.

There are three primary sources of periodicity in young stellar objects: stellar rotation, Keplerian rotation of an inner disk, and the orbital motion of a close binary companion. Stellar rotation can manifest itself via localized hot or cool spots, or via interactions between the stellar magnetic field and the inner disk. In the case of L54361, we reject stellar rotation effects on several grounds: (1) rotation periods for pre-main-sequence stars range from a few days to two weeks[18] (and protostars are typically faster still[19]), all lower than the measured period; (2) dark spots produce sinusoidal light curves, with amplitudes of a few tenths of a magnitude in the optical and declining to longer wavelengths; (3) hot spots tend to produce less obvious periodicity owing to the more stochastic nature of accretion, and can exhibit phase-dependent asymmetric illumination of the circumstellar material as they rotate with the star[20], which we do not see.

Regarding phenomena related to Keplerian rotation of an inner disk, persistent asymmetric structures such as warps in the inner disk can produce periodic obscuration of both single stars[7] and binary systems[21,22]. We argue that the data are not consistent with this scenario in several respects: (1) obscuration localized to the disk plane would not affect light propagating in the perpendicular direction through the outflow cavity; (2) obscuration events produce characteristic light curve 'dips', whereas we see a positive pulse-like shape; (3) as we show below, the MIR and far-infrared flux of L54361 originates

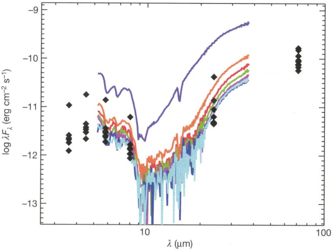


Figure 1 | Multi-epoch spectral energy distribution of L54361. Our complete set of observations taken during cryogenic Spitzer operations are shown, including photometry from all four IRAC channels at 3.6–8 μm and the MIPS 24 and 70 μm channels (diamonds), as well as 7 epochs of IRS spectroscopy (in chronological order: red, green, magenta, blue, cyan, purple and orange lines). Each single-epoch SED exhibits a shape characteristic of class I objects, with the flux rising sharply to longer wavelengths and a strong silicate absorption feature at 8–12 μm. Between epochs, however, the flux varies by as much as an order of magnitude at all wavelengths, with only a slightly shallower spectral slope in the ~15–70-μm continuum as the flux increases.

---

[1]Space Telescope Science Institute, 3700 San Martin Drive, Baltimore, Maryland 21218, USA. [2]National Optical Astronomy Observatory, Tucson, Arizona 85719, USA. [3]Steward Observatory, 933 North Cherry Avenue, University of Arizona, Tucson, Arizona 85721, USA. [4]Max-Planck-Institut fur Astronomie, Konigstuhl 17, 69117 Heidelberg, Germany. [5]Department of Astronomy, University of Massachusetts, Amherst, Massachusetts 01003, USA.

mostly in the infalling envelope, whose total flux would not be significantly affected by localized stellar obscuration.

The third possibility, a connection to binary motion, is plausible in terms of the length of the period of L54361, although we do not yet have direct evidence of a companion. The pulsed accretion scenario could explain both the shape and amplitude of the light curve. Circumbinary disk simulations consistently show gap-clearing by gravitational torques, followed by accretion streams that feed material onto the central stars[1–3]. For certain binary architectures, particularly in the case of a highly eccentric orbit, the stellar accretion depends on orbital phase, with the highest accretion rates typically associated with periastron passages. A qualitatively similar process has also been suggested for some X-ray binaries, at least one of which has exhibited optical light curves similar to the MIR behaviour of L54361[23,24].

An increase in the accretion luminosity as a result of the binary interaction increases the irradiation heating of circumstellar dust, which then reradiates the energy in the MIR where we observe it. To help test this hypothesis against our observations, we calculated radiative transfer models of protostellar dust emission and scattering[25]. The models include three components that are typical of protostellar systems: infalling envelope, accretion disk and central star. Holding all parameters fixed except for the accretion luminosity, we are able to match the change between SEDs corresponding to two different pulse phases (Fig. 4). The models show a relatively weak wavelength dependence as a function of luminosity, with a slightly flatter spectral slope at about 15 to 70 μm at higher luminosity as a result of optical depth effects in the envelope, in relatively good agreement with the observations.

We do not yet have any direct measure of the central object or its multiplicity status. As our models show, however, the bolometric luminosity provides an estimate of the stellar plus accretion luminosity. Assuming that the low end of the measured range is representative of the stellar luminosity, the combined stellar mass can be roughly estimated by comparing to a theoretical protostellar birthline[26] on an Hertzsprung–Russell diagram. We derive a value of ~0.2 solar masses ($M_\odot$; probably an upper limit because some contribution from accretion is likely, although the luminosity may also be somewhat underestimated because of scattering). Conversely, assuming that the upper end of the range of measured luminosity is due entirely to accretion luminosity and adopting the above stellar mass, we derive a maximum mass accretion rate of $10^{-6} M_\odot \mathrm{yr}^{-1}$. This is at the upper end of the range of values measured from standard accretion diagnostics[27]. Spectroscopic observations are needed to verify an accretion signature, as well as characterize the binary orbit.

Why L54361 exhibits such a strong and regular signature, unlike the T Tauri-type pulsed accretors observed previously, remains unknown. There may be a connection to its earlier evolutionary stage, in which the infalling envelope provides a steady supply of material to the circumbinary disk. By contrast, T Tauri binaries are older by about a factor of ten, have long since dissipated their natal envelopes, and accrete at lower mean rates. Perhaps stochastic variability from other

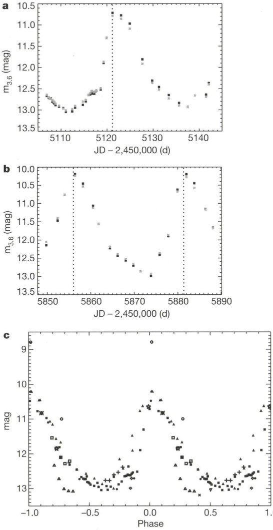


Figure 2 | Photometric light curves for L54361. **a**, **b**, IRAC 3.6 μm (black) and 4.5 μm (grey, scaled down to match 3.6) magnitudes from the autumn 2009 (**a**) and autumn 2011 (**b**) Spitzer observing campaigns. Note that the 3σ photometric uncertainties are equal to or smaller than the symbol size. The dashed line in **a** marks the observed peak time, which we set as the fiducial epoch for phase $= 0$. The dashed lines in **b** mark the predicted peak times assuming a periodicity of 25.34 d. **c**, The phased photometric light curve of L54361, assuming a period of 25.34 d and the phase-zero epoch Julian date (JD) 2,455,121.203. Included are measurements taken at three separate wavelengths. Each symbol type represents a contiguous set of photometry: cryo-Spitzer IRAC 3.6 μm (plus signs, asterisk, cross), warm Spitzer IRAC 3.6 μm (filled squares, filled triangles), MIPS 24 μm (inverted triangle, open diamonds, open stars, open triangles), IRS 24 μm (open circles, open squares) and HST WFC3 1.6 μm (filled circles). The 24 μm photometry values are offset by $+7.3$ mag and WFC3 photometry offset by -5.3 mag to place everything on the same scale.

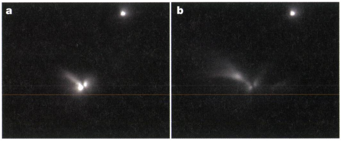


Figure 3 | Near-infrared images of L54361. **a**, **b**, Portions of images taken with HST/WFC3 at 1.6 μm at two epochs corresponding to pulse phases of 0 (**a**) and 0.3 (**b**). North is up and east is to the left. L54361 is the extended source just below the centre of the images; the point source at upper right is another young stellar object, LRLL 1843. The light from L54361 subtends roughly 14″ (~4,000 AU at the distance of the IC 348 region) in **a**, and about 50″ (~15,000 AU) in **b**. Most if not all of this light is probably the result of scattering off circumstellar dust in the protostellar envelope. An apparent edge-on disk is visible at the centre of the object, and three separate structures indicative of outflow cavities extend to the northwest, southwest and northeast. The extent and morphology of the scattered light changes substantially between epochs as a result of the propagation of the pulse peak light. (See Supplementary Information for the complete set of HST images.)

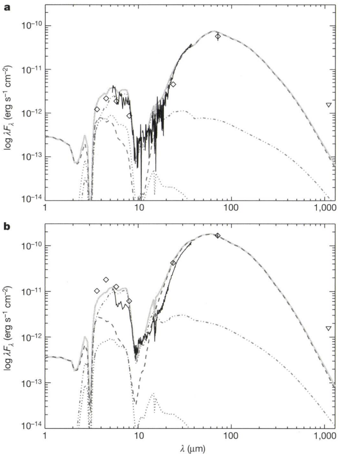


Figure 4 | Protostellar spectral energy distribution models for L54361. **a**, **b**, Comparison of models to two sets of observations roughly representing pulse phases of 0.4 (**a**) and 0.15 (**b**). The photometry and spectroscopy are not simultaneous but were selected to correspond roughly to a common flux level. Dashed lines indicate scattering and emission from the infalling envelope; dashed-double-dotted lines represent scattering and emission from the circumstellar accretion disk; dotted lines represent reddened flux from the central star; the solid grey line shows the total flux from all these components. The best-fit parameters include an envelope infall rate of $3 \times 10^{-6} M_\odot\,\mathrm{yr}^{-1}$ (assuming a stellar mass of $0.2 M_\odot$), envelope centrifugal radius $R_c = 30$ AU, outflow cavity opening angle $\theta = 30°$, inclination angle of the outflow/stellar rotation axis to the line of sight $i = 70°$, total central luminosity (stellar plus accretion) $L = 0.5 L_\odot$ (**a**) and $1.3 L_\odot$ (**b**), and the fraction of L due to the stellar luminosity $\eta_{\mathrm{star}} = 0.5$ (**a**) and 0.19 (**b**). The only parameter that was actually changed between the two models was η_{star}, with appropriate values so that the stellar luminosity remained constant while the accretion luminosity changed by a factor of $\sim$4. (See Supplementary Information for details of the model calculations and parameters.)

sources such as stellar magnetic interactions or disk turbulence can overwhelm the periodic signature in older stars. It is also possible that the particular orbital parameters of L54361 are rare but more favourable for modulating the accretion flow, such as a very large eccentricity.

Received 21 February; accepted 31 October 2012.

Published online 2 January 2013.

1. Artymowicz, P. & Lubow, S. H. Dynamics of binary-disk interaction. 1: Resonances and disk gap sizes. *Astrophys. J.* **421,** 651–667 (1994).
2. Artymowicz, P. & Lubow, S. H. Mass flow through gaps in circumbinary disks. *Astrophys. J.* **467,** L77–L80 (1996).
3. Gunther, R. & Kley, W. Circumbinary disk evolution. *Astron. Astrophys.* **387,** 550–559 (2002).
4. Shi, J., Krolik, J. H., Lubow, S. H. & Hawley, J. F. Three dimensional MHD simulation of circumbinary accretion disks: disk structures and angular momentum transport. *Astrophys. J.* **749,** 118–144 (2012).
5. Quintana, E. V. & Lissauer, J. J. Terrestrial planet formation surrounding close binary stars. *Icarus* **185,** 1–20 (2006).
6. Paardekooper, S.-J., Thebault, P. & Mellema, G. Planetesimal and gas dynamics in binaries. *Mon. Not. R. Astron. Soc.* **386,** 973–988 (2008).
7. Bouvier, J. *et al.* Magnetospheric accretion-ejection processes in the classical T Tauri star AA Tauri. *Astron. Astrophys.* **463,** 1017–1028 (2007).
8. Nguyen, D. C. & Scholz, A. van Kerkwijk, M. H., Jayawardhana, R. & Brandeker, A. How variable is accretion in young stars? *Astrophys. J.* **694,** L153–L157 (2009).
9. Flaherty, K. *et al.* Infrared variability of evolved protoplanetary disks: evidence for scale height variations in the inner disk. *Astrophys. J.* **748,** 71–100 (2012).
10. Morales-Calderón, M. *et al.* Ysovar: the first sensitive, wide-area, mid-infrared photometric monitoring of the Orion nebula cluster. *Astrophys. J.* **733,** 50–59 (2011).
11. Liu, M. C. *et al.* Mid-infrared imaging of young stellar objects. *Astrophys. J.* **461,** 334–344 (1996).
12. Barsony, M., Ressler, M. E. & Marsh, K. A. A. Mid-infrared imaging survey of embedded young stellar objects in the ρ Ophiuchi cloud core. *Astrophys. J.* **630,** 381–399 (2005).
13. Kóspál, A. *et al.* The outburst and nature of two young eruptive stars in the North America/Pelican Nebula Complex. *Astron. Astrophys.* **527,** A133 (2011).
14. Basri, G., Johns-Krull, C. M. & Mathieu, R. D. The classical T Tauri spectroscopic binary DQ tau. II. Emission line variations with orbital phase. *Astron. J.* **114,** 781–792 (1997).
15. Mathieu, R. D. *et al.* The classical T Tauri spectroscopic binary DQ tau. I. Orbital elements and light curves. *Astron. J.* **113,** 1841–1854 (1997).
16. Jensen, E. L. N. *et al.* Periodic accretion from a circumbinary disk in the young binary UZ Tau E. *Astron. J.* **134,** 241–251 (2007).
17. Stark, D. P., Whitney, B. A., Stassun, K. & Wood, K. Near-infrared synthetic images of protostellar disks and envelopes. *Astrophys. J.* **649,** 900–913 (2006).
18. Nordhagen, S., Herbst, W., Rhode, K. L. & Williams, E. C. The variability and rotation of pre-main-sequence stars in IC 348: does intracluster environment influence stellar rotation? *Astron. J.* **132,** 1555–1570 (2006).
19. Covey, K. R., Greene, T. P., Doppmann, G. W. & Lada, C. J. The angular momentum content and evolution of class I and flat-spectrum protostars. *Astron. J.* **129,** 2765–2776 (2005).
20. Wood, K. & Whitney, B. Scattered light signatures of magnetic accretion in classical T Tauri stars. *Astrophys. J.* **506,** L43–L45 (1998).
21. Herbst, W. *et al.* The light curve of the weakly accreting T Tauri binary KH 15D from 2005–2010: insights into the nature of its protoplanetary disk. *Astron. J.* **140,** 2025–2035 (2010).
22. Plavchan, P., Gee, A. H., Stapelfeldt, K. & Becker, A. The peculiar periodic YSO WL 4 in ρ Ophiuchius. *Astrophys. J.* **684,** L37–L40 (2008).
23. Skinner, G. K. Observations of optical flares in the recurrent X-ray transient A0538–66. *Nature* **288,** 141–143 (1980).
24. Densham, R. H., Charles, P. A., Menzies, J. W., van der Klis, M. & van Paradijs, J. Four outburst cycles of A0538–66: evidence for a rapidly evolving envelope around the primary. *Mon. Not. R. Astron. Soc.* **205,** 1117–1133 (1983).
25. Kenyon, S. J., Calvet, N. & Hartmann, L. The embedded young stars in the Taurus-Auriga molecular cloud. I — Models for spectral energy distributions. *Astrophys. J.* **414,** 676–694 (1993).
26. Hartmann, L., Cassen, P. & Kenyon, S. J. Disk accretion and the stellar birthline. *Astrophys. J.* **475,** 770–785 (1997).
27. White, R. J. & Hillenbrand, L. A. On the evolutionary status of class I stars and Herbig-Haro energy sources in Taurus-Auriga. *Astrophys. J.* **616,** 998–1032 (2004).

Supplementary Information is available in the online version of the paper.

Acknowledgements This work was supported in part by NASA through Spitzer and HST GO contracts. We thank S. Lubow, M. Livio and N. Calvet for discussions. E.F. was visiting the Infrared Processing and Analysis Center, Caltech, during the course of this work.

Author Contributions J.M. and K.F. designed the Spitzer observations. Z.B. and R.G. reduced the IRAC images and compiled the photometry, while J.M. reduced and analysed the MIPS data. J.M. and E.F. extracted and analysed the IRS spectroscopy. J.M. designed the HST observations and analysed the images. E.F. calculated the radiative transfer models and fitted the observed SEDs. All authors contributed to the writing of the paper.

Author Information Reprints and permissions information is available at www.nature.com/reprints. The authors declare no competing financial interests. Readers are welcome to comment on the online version of the paper. Correspondence and requests for materials should be addressed to J.M. (muzerol@stsci.edu).

LETTER

doi:10.1038/nature11780

Planetary system disruption by Galactic perturbations to wide binary stars

Nathan A. Kaib[1,2]†, Sean N. Raymond[3,4] & Martin Duncan[1]

Nearly half the exoplanets found within binary star systems reside[1] in very wide binaries with average stellar separations greater than 1,000 astronomical units (one astronomical unit (AU) being the Earth–Sun distance), yet the influence of such distant binary companions on planetary evolution remains largely unstudied. Unlike their tighter counterparts, the stellar orbits of wide binaries continually change under the influence of the Milky Way's tidal field and impulses from other passing stars. Here we report numerical simulations demonstrating that the variable nature of wide binary star orbits dramatically reshapes the planetary systems they host, typically billions of years after formation. Contrary to previous understanding[2], wide binary companions may often strongly perturb planetary systems, triggering planetary ejections and increasing the orbital eccentricities of surviving planets. Although hitherto not recognized, orbits of giant exoplanets within wide binaries are statistically more eccentric than those around isolated stars. Both eccentricity distributions are well reproduced when we assume that isolated stars and wide binaries host similar planetary systems whose outermost giant planets are scattered beyond about 10 AU from their parent stars by early internal instabilities. Consequently, our results suggest that although wide binaries eventually remove the most distant planets from many planetary systems, most isolated giant exoplanet systems harbour additional distant, still undetected planets.

Unlike binary stars with separations below $\sim 10^3$ AU, very widely separated binaries ('wide binaries') are only weakly bound by self-gravity, leaving them susceptible to outside perturbations. As a result, the Milky Way's tide and impulses from other passing stars strongly perturb wide-binary orbits[3,4]. These perturbations, which are fairly independent of the orbiting object's mass, are also known to dramatically affect the dynamics of Solar System comets at similar orbital distances[5,6]. Galactic perturbations drive a pseudo-random walk in the pericentres of these comets (the closest approach distances to the Sun)[5,7]. The same effect will occur in wide-binary orbits. Thus, even if a very wide binary's initial pericentre is quite large, it will inevitably become very small at some point if it remains gravitationally bound and evolves long enough. Such low-pericentre phases will produce close stellar passages between binary members, with potentially devastating consequences for planetary systems in these binaries[8,9]. Counterintuitively, we therefore suspect that wide binary companions could more dramatically affect planetary system evolution than tight binaries.

To investigate this scenario, we use the Mercury simulation package to perform 2,600 simulations modelling the orbital evolution of our Sun's four giant planets (on their current orbits) in the presence of a very wide binary companion[10]. These simulations are listed as set A in Table 1, which briefly summarizes the initial conditions of our different simulations (see Supplementary Information for details). An example simulation is shown in Fig. 1. Initially, the binary companion has no effect on the planets' dynamics because its starting pericentre (q; solid black line in Fig. 1) is $\sim$3,000 AU. However, after 1 Gyr of evolution, Galactic perturbations drive the binary pericentre near 100 AU, exciting the eccentricities of Neptune and Uranus. Once again, at 3.5 Gyr, the binary passes through another low-pericentre phase, this time triggering the ejection of Uranus. Last, at 7.2 Gyr, the binary makes a final excursion to low q, causing Neptune's ejection.

Such behaviour is not unusual. Depending on the binary's mass and semimajor axis (mean separation, or a_*), Fig. 2a demonstrates that $\sim$30–60% of planetary systems in simulation set A experience instabilities causing one or more planetary ejections after 10 Gyr (the approximate age of our Galaxy's thin disk). Even though binaries with smaller semimajor axes are less affected by Galactic perturbations, Fig. 2a shows that the influence of binary semimajor axis on planetary instability rates is weak. This is because tighter binaries make pericentre passages at a higher frequency. In addition, when they reach low-q phases they remain stuck there for a much longer time than wider binaries. As Fig. 2b shows, both of these effects cause tighter binaries to become lethal (that is, destabilizing) at a much larger pericentre, offsetting the Galaxy's diminished influence. Most binary-triggered instabilities are very delayed. For binaries with $a_* \gtrsim 2{,}000$ AU, Fig. 2c shows that well over 90% of instabilities occur after at least 100 Myr of evolution, well after planet formation is complete. For tighter binaries, many more begin in orbits that destabilize the planets nearly instantly.

Although planets are believed to form on nearly circular orbits[11], most known giant planets ($m \sin i > 1\,M_{\rm Jup}$, where m is the planet's mass, i its orbital inclination and $M_{\rm Jup}$ is Jupiter's mass) have significant non-zero orbital eccentricities (eccentricities of less massive planets are known to be lower, or 'colder')[12]. This observed eccentricity distribution can be reproduced remarkably well when systems of circularly orbiting planets undergo internal dynamical instabilities causing planet–planet scattering events that eject some planets and excite the survivors' eccentricities[13–16]. For planetary systems within wide binaries, Fig. 2a predicts that many should undergo additional

Table 1 | Initial conditions of simulation sets

Name of simulation set	Number of planets	Planet masses ($M_{\rm Jup}$)	Planet a-range (AU)	Binary mass ($M_\odot$)	Binary a^* (AU)	External perturbations included
A	4	SS	SS	0.1–1.0	1,000–30,000	Tide + stars
B1	3	0.5 to $\sim$15	2 to $\sim$15	None	None	None
B2	3	0.5 to $\sim$15	2 to $\sim$15	0.4	1,000–30,000	Tide + stars
B3	3	0.5 to $\sim$15	2 to $\sim$15	0.4	1,000–30,000	None

SS refers to planetary systems resembling the Solar System's four giant planets. Tide + stars refers to perturbations from the Galactic tide and passing field stars. $M_{\rm Jup}$, mass of Jupiter; a, semimajor axis; $M_\odot$, solar mass.

[1] Department of Physics, Queen's University, Kingston, Ontario K7L 3N6, Canada. [2] Canadian Institute for Theoretical Astrophysics, University of Toronto, Toronto, Ontario M5S 3H8, Canada. [3] Université de Bordeaux, Observatoire Aquitain des Sciences de l'Univers, 2 rue de l'Observatoire, BP 89, F-33271 Floirac Cedex, France. [4] CNRS, UMR 5804, Laboratoire d'Astrophysique de Bordeaux, 2 rue de l'Observatoire, BP 89, F-33271 Floirac Cedex, France. †Present address: Center for Interdisciplinary Exploration and Research in Astrophysics and Department of Physics and Astronomy, Northwestern University, 2131 Tech Drive, Evanston, Illinois 60208, USA.

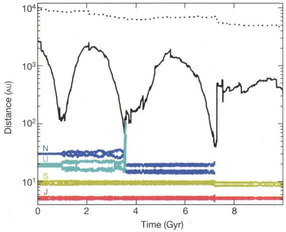


Figure 1 | Typical binary-triggered disruption. Simulation of a binary-triggered instability in a planetary system resembling our own Solar System. The pericentre and apocentre are plotted for Jupiter (J; red lines), Saturn (S; gold), Uranus (U; cyan) and Neptune (N; blue). The binary's semimajor axis (dotted black line) and pericentre (solid black line) are also shown.

dynamical instabilities triggered by their stellar companions. Thus, these systems should experience an even greater number of planet–planet scattering events than isolated planetary systems.

This raises the possibility that the eccentricities of exoplanets may hold a signature of the dynamical process illustrated in Fig. 1. Indeed, the overall distribution of exoplanet eccentricities provides compelling evidence of our disruptive mechanism. Figure 3a compares the observed eccentricity distribution of all Jovian-mass ($m \sin i > 1\,M_{\mathrm{Jup}}$) exoplanets found in binaries[1] with the distribution of Jovian-mass planets around isolated stars. As can be seen, the distribution of planets within wide binaries is significantly hotter (or more biased to higher eccentricities) than planetary systems without known stellar companions. A Kolmogorov-Smirnov test returns a probability (*P*-value) of only 0.6% that such a poor match between the two data sets will occur if they sample the same underlying distribution. Thus, we reject the null hypothesis that the distributions are the same. Although it consists of just 20 planets, our wide-binary planetary sample contains the two most eccentric known exoplanet orbits, HD 80606b and HD 20782b (see Fig. 2). Furthermore, these excited eccentricities seem to be confined to only very wide binary systems. Figure 3a also shows the eccentricity distribution of planets residing in binaries with average separations below 10^3 AU. Unlike the case for wider binaries, here we see that these eccentricities match very closely with the isolated distribution. (A Kolmogorov-Smirnov test returns a *P*-value of 91%.) This suggests that the variable nature of distant binary orbits is crucial to exciting planetary orbits. Large eccentricities of planets within binaries have previously been explained with the Kozai resonance[17–20], yet this effect should be most evident in these tighter binary systems.

We perform additional simulations, attempting to explain the observed eccentricity excitation in Fig. 3a with the mechanism illustrated in Fig. 1. These additional simulation sets are summarized in Table 1 (B1–B3). Unlike the internally stable planetary systems in simulation set A, these simulated systems consist of three approximately Jovian-mass planets started in unstable configurations (to induce planet–planet scattering) and evolved for 10 Gyr (see Supplementary Information). In the simulation sets presented in Fig. 3b, we naturally reproduce both observed eccentricity distributions using the same initial planetary systems. When our planetary systems are run in isolation (set B1 in Table 1) planet–planet scattering caused by internal instabilities yields the observed planetary eccentricities for isolated stars (Kolmogorov-Smirnov test *P*-value of 42%). Then when a $0.4M_{\odot}$ binary companion is added to each system (set B2 in Table 1) the eccentricity distribution is heated further, and again the match to

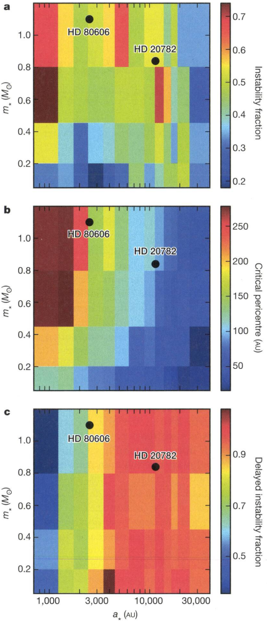


Figure 2 | Disruption as a function of binary mass and separation. **a–c**, Maps of the fraction of systems in set A (see Table 1) that lost at least one planet via instability (**a**), the median binary pericentre below which an instability is induced in planetary systems (**b**) and the fraction of instabilities that occur after the first 100 Myr of evolution (**c**). In each panel, binary mass (m_*) is plotted on the *y* axis, and the binary semimajor axis (a_*) is plotted on the x axis; black data points mark the masses and presumed semimajor axes of the HD 80606 and HD 20782 binaries, which host the two most eccentric known planetary orbits[26,27]. Although HD 80606b has been reproduced with a Kozai-driven mechanism, this process is markedly slower in even wider binaries such as HD 20782b[28]. Moreover, the presence of more than one planet suppresses these Kozai oscillations[28–30]. However, our disruptive mechanism naturally collapses many systems to one planet, still enabling Kozai resonances to contribute to eccentricity excitation. Panel **a** suggests that binary-triggered instability rates become extremely high as binary semimajor axes drop below $\sim 10^3$ AU, which could mean that tighter binaries trigger planetary system instabilities even more efficiently than those plotted here. However, the initial conditions assumed for both our planetary orbits (Solar System-like) and binary orbits (isotropic) become questionable for binary semimajor axes below $\sim 10^3$ AU (see Supplementary Information). Another interesting aspect not immediately obvious in **c** is that instability times decrease at the largest binary semimajor axes. This is because such binaries are rapidly unbound (or 'ionized') by stellar impulses, making it impossible for these binaries to trigger instabilities at very late epochs (see Supplementary Information).

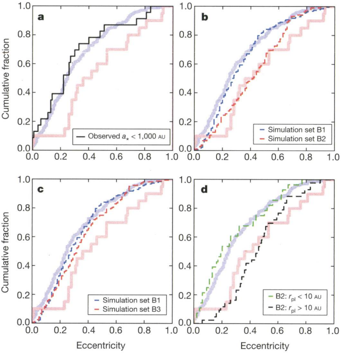


Figure 3 | Eccentricity excitation of planets of wide binaries. Comparison of planetary eccentricity distributions. In all panels, pale red lines correspond to observed systems with very wide ($a_* > 10^3$ AU) companions and pale blue lines are for observed isolated systems. Dashed lines describe the innermost planets in our simulated systems (sets B1–B3 in Table 1). In all distributions, planets with $a < 0.1$ AU are excluded to remove tidally circularized orbits. **a**, A comparison of observed exoplanet eccentricities within tighter ($a_* < 10^3$ AU) binaries (solid black line) to those observed in very wide ($a_* > 10^3$ AU) binaries and isolated systems. **b**, Eccentricities of simulated three-planet systems after 10 Gyr of evolution (sets B1 and B2 in Table 1). **c**, Simulations from **b** rerun with no Galactic perturbations (set B3 in Table 1). **d**, The final eccentricities of two different subgroups of wide binary simulations from **b**: systems that consisted of two planets extending beyond 10 AU at 10 Myr (black line), and two-planet systems confined inside 10 AU at 10 Myr (green). Notice in **b** that the presence of a wide binary does not seem to enhance the production of very extreme planetary eccentricities. However, one-quarter of our wide binary systems have planets driven into the central star (1.7 times the rate within isolated systems). Tidal dissipation not included in our models could strand these planets in very eccentric orbits before they collide with the central star[19,20]. Interestingly, binaries also completely strip 20% of our systems of planets, yielding naked stars that once hosted planets (see Supplementary Information).

observations is quite good, with a Kolmogorov-Smirnov test P-value of 46%.

In Fig. 3c, the match to observed planetary eccentricities is much poorer. Here we rerun our binary simulations with Galactic perturbations shut off to yield static binary orbits (set B3 in Table 1). In this case, the eccentricity distribution is barely more excited than the isolated cases, indicating that the variable nature of wide-binary orbits is crucial to heating planetary eccentricities. Otherwise, most stellar companions always remain far from the planets.

In Fig. 3d we reexamine simulation set B2 to determine which types of planetary systems are most influenced by wide binary companions. By examining the planetary systems after only 10 Myr, we can view them after most have experienced internal instabilities but before the binary has played a large role (because its effects are delayed). We find that 70% of our planetary systems have collapsed to two planets. (The remaining systems are composed of nearly equal numbers of one-planet and three-planet systems.) We then split these two-planet systems into those with the outer planet beyond 10 AU and those with both planets confined inside 10 AU. In Fig. 3d, the final ($t = 10$ Gyr) eccentricity distribution is shown for both subgroups of planetary systems. As can be seen, the more extended planetary systems eventually yield much more excited eccentricities compared to the compact systems. This is because binaries do not have to evolve to such low pericentres to disrupt extended systems. In fact, the observed wide-binary planetary eccentricity distribution cannot be matched without using wide binaries with planets beyond 10 AU ($P = 1.6\%$ from a Kolmogorov-Smirnov test). Assuming planets form similarly in wide binaries and isolated systems, the planetary eccentricity excitation observed within wide binaries may offer new constraints on the bulk properties of isolated giant exoplanet systems, which dominate the giant exoplanet catalogue. Whereas most detection efforts are currently insensitive to planets with periods beyond ~10 yr, our work suggests that massive, longer-period planets (beyond ~10 AU) should be common around isolated stars. Indeed, such distant planets have recently been directly observed[21] and microlensing results suggest many such planets reside far from host stars[22].

Owing to the variable nature of their orbits, very distant binary companions may affect planetary evolution at least as strongly as their tighter counterparts. This represents a paradigm shift in our understanding of planet-hosting binaries, because previous work tended to assume that only tighter binaries strongly influence planetary system evolution[2,23]. Intriguingly, the eccentricities of planets in wide binaries may provide new constraints on the intrinsic architectures of all planetary systems. To develop this prospect further, searches for common proper motion companions to planet-hosting stars should be continued and expanded[2,23–25].

Received 27 July; accepted 1 November 2012.
Published online 6 January 2013.

1. Roell, T., Neuhäuser, R., Seifahrt, A. & Mugrauer, M. Extrasolar planets in stellar multiple systems. *Astron. Astrophys.* **542,** A92 (2012).
2. Desidera, S. & Barbieri, M. Properties of planets in binary systems. The role of binary separation. *Astron. Astrophys.* **462,** 345–353 (2007).
3. Heggie, D. C. & Rasio, F. A. The effect of encounters on the eccentricity of binaries in clusters. *Mon. Not. R. Astron. Soc.* **282,** 1064–1084 (1996).
4. Jiang, Y.-F. & Tremaine, S. The evolution of wide binary stars. *Mon. Not. R. Astron. Soc.* **401,** 977–994 (2010).
5. Oort, J. H. The structure of the cloud of comets surrounding the Solar System and a hypothesis concerning its origin. *Bull. Astron. Inst. Neth.* **11,** 91–110 (1950).
6. Heisler, J. & Tremaine, S. The influence of the galactic tidal field on the Oort comet cloud. *Icarus* **65,** 13–26 (1986).
7. Kaib, N. A. & Quinn, T. Reassessing the source of long-period comets. *Science* **325,** 1234–1236 (2009).
8. Adams, F. C. & Laughlin, G. Constraints on the birth aggregate of the Solar System. *Icarus* **150,** 151–162 (2001).
9. Zakamska, N. L. & Tremaine, S. Excitation and propagation of eccentricity disturbances in planetary systems. *Astron. J.* **128,** 869–877 (2004).
10. Chambers, J. E., Quintana, E. V., Duncan, M. J. & Lissauer, J. J. Symplectic integrator algorithms for modeling planetary accretion in binary star systems. *Astron. J.* **123,** 2884–2894 (2002).
11. Lissauer, J. J. Planet formation. *Annu. Rev. Astron. Astrophys.* **31,** 129–172 (1993).
12. Wright, J. T. *et al.* Ten new and updated multiplanet systems and a survey of exoplanetary systems. *Astrophys. J.* **693,** 1084–1099 (2009).
13. Jurić, M. & Tremaine, S. Dynamical origin of extrasolar planet eccentricity distribution. *Astrophys. J.* **686,** 603–620 (2008).
14. Ford, E. B. & Rasio, F. A. Origins of eccentric extrasolar planets: testing the planet-planet scattering model. *Astrophys. J.* **686,** 621–636 (2008).
15. Malmberg, D. & Davies, M. B. On the origin of eccentricities among extrasolar planets. *Mon. Not. R. Astron. Soc.* **394,** L26–L30 (2009).
16. Raymond, S. N., Armitage, P. J. & Gorelick, N. Planet-planet scattering in planetesimal disks. II. Predictions for outer extrasolar planetary systems. *Astrophys. J.* **711,** 772–795 (2010).
17. Kozai, Y. Secular perturbations of asteroids with high inclination and eccentricity. *Astron. J.* **67,** 591–598 (1962).
18. Holman, M., Touma, J. & Tremaine, S. Chaotic variations in the eccentricity of the planet orbiting 16 Cygni B. *Nature* **386,** 254–256 (1997).
19. Wu, Y. & Murray, N. Planet migration and binary companions: the case of HD 80606b. *Astrophys. J.* **589,** 605–614 (2003).
20. Fabrycky, D. & Tremaine, S. Shrinking binary and planetary orbits by Kozai cycles with tidal friction. *Astrophys. J.* **669,** 1298–1315 (2007).
21. Marois, C., Zuckerman, B., Konopacky, Q. M., Macintosh, B. & Barman, T. Images of a fourth planet orbiting HR 8799. *Nature* **468,** 1080–1083 (2010).
22. The Microlensing Observations in Astrophysics (MOA) Collaboration & The Optical Gravitational Lensing Experiment (OGLE) Collaboration. Unbound or distant planetary mass population detected by gravitational microlensing. *Nature* **473,** 349–352 (2011).

23. Eggenberger, A. *et al.* The impact of stellar duplicity on planet occurrence and properties. I. Observational results of a VLT/NACO search for stellar companions to 130 nearby stars with and without planets. *Astron. Astrophys.* **474,** 273–291 (2007).
24. Raghavan, D. *et al.* Two suns in the sky: stellar multiplicity in exoplanet systems. *Astrophys. J.* **646,** 523–542 (2006).
25. Mugrauer, M. *et al.* A search for wide visual companions of exoplanet host stars: the Calar Alto survey. *Astron. Nachr.* **327,** 321–327 (2006).
26. Jones, H. R. A. *et al.* High-eccentricity planets from the Anglo-Australian Planet Search. *Mon. Not. R. Astron. Soc.* **369,** 249–256 (2006).
27. Naef, D. *et al.* HD 80606 b, a planet on an extremely elongated orbit. *Astron. Astrophys.* **375,** L27–L30 (2001).
28. Innanen, K. A., Zheng, J. Q., Mikkola, S. & Valtonen, M. J. The Kozai mechanism and the stability of planetary orbits in binary star systems. *Astron. J.* **113,** 1915–1919 (1997).
29. Batygin, K., Morbidelli, A. & Tsiganis, K. Formation and evolution of planetary systems in presence of highly inclined stellar perturbers. *Astron. Astrophys.* **533,** A7 (2011).
30. Kaib, N. A., Raymond, S. N. & Duncan, M. J. 55 Cancri: a coplanar planetary system that is likely misaligned with its star. *Astrophys. J.* **742,** L24 (2011).

Supplementary Information is available in the online version of the paper.

Acknowledgements We thank J. Chambers and R. Roškar for discussions. This work was funded by a CITA National Fellowship and Canada's NSERC. S.N.R. thanks the PNP programme of CNRS and the NASA Astrobiology Institute's Virtual Planetary Laboratory team. Our computing was performed on the SciNet General Purpose Cluster at the University of Toronto.

Author Contributions N.A.K. performed the simulations and analysis and was the primary writer of this paper. S.N.R. and M.D. helped initiate the project and advised on simulations and analysis.

Author Information Reprints and permissions information is available at www.nature.com/reprints. The authors declare no competing financial interests. Readers are welcome to comment on the online version of the paper. Correspondence and requests for materials should be addressed to N.A.K. (nkaib@astro.queensu.ca).

LETTER

doi:10.1038/nature11728

Ultrahard nanotwinned cubic boron nitride

Yongjun Tian[1], Bo Xu[1], Dongli Yu[1], Yanming Ma[2], Yanbin Wang[3], Yingbing Jiang[4], Wentao Hu[1], Chengchun Tang[5], Yufei Gao[1], Kun Luo[1], Zhisheng Zhao[1], Li-Min Wang[1], Bin Wen[1], Julong He[1] & Zhongyuan Liu[1]

Cubic boron nitride (cBN) is a well known superhard material that has a wide range of industrial applications. Nanostructuring of cBN is an effective way to improve its hardness by virtue of the Hall–Petch effect—the tendency for hardness to increase with decreasing grain size[1,2]. Polycrystalline cBN materials are often synthesized by using the martensitic transformation of a graphite-like BN precursor, in which high pressures and temperatures lead to puckering of the BN layers[3]. Such approaches have led to synthetic polycrystalline cBN having grain sizes as small as ∼14 nm (refs 1, 2, 4, 5). Here we report the formation of cBN with a nanostructure dominated by fine twin domains of average thickness ∼3.8 nm. This nanotwinned cBN was synthesized from specially prepared BN precursor nanoparticles possessing onion-like nested structures with intrinsically puckered BN layers and numerous stacking faults. The resulting nanotwinned cBN bulk samples are optically transparent with a striking combination of physical properties: an extremely high Vickers hardness (exceeding 100 GPa, the optimal hardness of synthetic diamond), a high oxidization temperature (∼1,294 °C) and a large fracture toughness (>12 MPa m$^{1/2}$, well beyond the toughness of commercial cemented tungsten carbide, ∼10 MPa m$^{1/2}$). We show that hardening of cBN is continuous with decreasing twin thickness down to the smallest sizes investigated, contrasting with the expected reverse Hall–Petch effect below a critical grain size or the twin thickness of ∼10–15 nm found in metals and alloys.

Polycrystalline cubic boron nitride possesses high hardness and high chemical stability and is the best-known material for cutting ferrous and carbide-forming hard substances where diamond completely fails. However, commercial polycrystalline cBNs have grain sizes in the micrometre scale and a Vickers hardness (H_V) of only 33–45 GPa (refs 6, 7), much lower than that of diamond (∼100 GPa)[8]. The application of cBN is therefore limited. For decades, scientists have designed various ways to increase the hardness of polycrystalline cBN[1,2,4,5] while retaining its excellent chemical stability.

The Hall–Petch effect[9,10], through which material hardness increases with decreasing grain size, has been widely observed. For polycrystalline cBN or diamond, reducing the grain size has been the most widely used approach for enhancing hardness. Conditions of high pressure and high temperature (HPHT) are required for the synthesis because temperature enhances the nucleation rate, and pressure suppresses grain growth by reducing the atomic diffusion responsible for crystallization[2,11]. Experimentally, nanograined (ng) diamonds with grain sizes of 10–30 nm have been synthesized[11]. High Knoop hardnesses of up to 110–140 GPa have been reported. Efforts in reducing the grain size further to below 10 nm have been unsuccessful as a result of intergranular fractures in these bulk ng-diamonds[12]. For cBN, the best attempts[1,2] performed so far have produced ng-cBN samples at pressures ∼20 GPa with a minimum grain size of 14 nm and a record hardness of 85 GPa. This grain size is close to the critical value at which the Hall–Petch relation breaks down, as demonstrated in metals and alloys[13–16]. It remains a technical challenge to synthesize polycrystalline cBNs with perfect grain boundaries and smaller grain sizes, because numerous high-energy grain boundaries in a polycrystalline material generate large driving forces for grain growth.

Fine twin domains can be stabilized within materials capable of twinning. Coherent twin boundaries possess excess energy that is typically about one order of magnitude lower than that of grain boundaries[15]. This makes nanotwinned structures energetically more stable than their nanograined counterparts, thereby allowing finer control of domain size. Twin boundaries have been experimentally verified to have a hardening effect identical to that of high-angle grain boundaries[15], because, similarly to grain boundaries, twin boundaries also serve as barriers against gliding dislocations. Hence, the synthesis of nanotwinned substructures within grains may be a feasible route to achieving ultrafine microstructural sizes approaching a few nanometres. In practice, however, a strategically designed experimental protocol essential to fine-tune nanotwin substructures has yet to be achieved.

The use of graphite-like BN as precursors merely generates ng-cBN[1,2,4,5], although twin domains are occasionally observed during the growth of cBN single crystals as undesired defects[17,18]. Here we report the synthesis of ubiquitously nanotwinned cBN (nt-cBN) with an average twin thickness of 3.8 nm under HPHT conditions, using a special turbostratic onion-like BN (oBN) nanoparticle precursor material.

The starting oBN nanoparticles (∼30–150 nm in diameter) consisted of concentric BN spherical shells (Fig. 1a, b) with abundant puckering and stacking faults. The inter-shell spacing was determined by X-ray diffraction at 0.351 nm (Supplementary Fig. 2a), which is slightly larger than the inter-layer spacings in hexagonal (0.3328 nm) and rhombohedral (0.334 nm) BN[17]. The oBN precursors were pressed into a pellet (2.5 mm in diameter and 1 mm in height) and subjected to HPHT treatments with a multi-anvil apparatus. X-ray diffraction and Raman spectroscopy were performed on the recovered samples to characterize structural phase transformations (Supplementary Fig. 2). At 15 GPa, no transformation was observed at temperatures below 1,000 °C. Transformation into a translucent cubic phase mixed with a small amount of the wurtzite phase occurred above 1,200 °C. The sample transformed entirely into transparent cBN (Fig. 1c and Supplementary Fig. 1) above 1,600 °C. Other synthetic trials performed at different pressures from 12 GPa to 25 GPa showed a similar temperature dependence of the phase transformation. For comparison, a considerably higher temperature (2,200 °C) was used to acquire pure cBN when graphite-like BN precursors were employed[1]. The use of the oBN precursor is responsible for the significantly reduced synthetic temperature (by about 600 °C) and the stabilization of nanotwinned microstructure (see below).

The synthetic nt-cBN bulk samples had a diameter of about 2 mm and possessed a unique homogeneous microstructure. Typical transmission electron microscopy (TEM) and high-resolution TEM (HRTEM) images of an as-synthesized bulk sample are shown in Fig. 1d–f. Grains were about the same size as the original onions, irregular in shape but approximately equiaxed. Selected area electron diffraction (inset of Fig. 1d) confirmed that the sample was pure cBN containing randomly oriented nanograins. HRTEM revealed a high

---

[1]State Key Laboratory of Metastable Materials Science and Technology, Yanshan University, Qinhuangdao 066004, China. [2]State Key Laboratory for Superhard Materials, Jilin University, Changchun 130012, China. [3]Center for Advanced Radiation Sources, University of Chicago, Chicago, Illinois 60439, USA. [4]TEM Laboratory, University of New Mexico, Albuquerque, New Mexico 87131, USA. [5]School of Material Science and Engineering, Hebei University of Technology, Tianjin 300130, China.

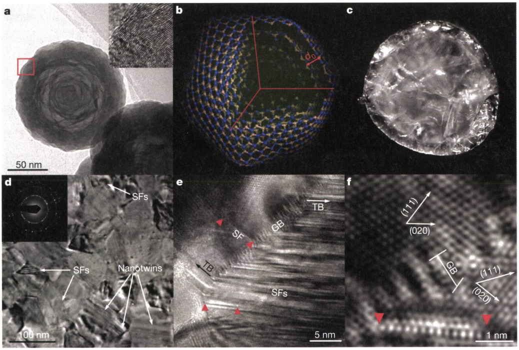


Figure 1 | Starting oBN nanoparticles and nt-cBN bulk synthesized at 15 GPa and 1,800 °C. **a**, TEM image of oBN nanoparticles. Inset, HRTEM image corresponding to the position marked with the red box, showing defects including lattice puckering, bending and stacking faults. **b**, Schematic icosahedral model of a five-shell oBN nanoparticle. d_0 is the inter-shell spacing. **c**, Photograph of an nt-cBN bulk sample with a diameter of about 2 mm. **d**, TEM image of a typical microstructure in nt-cBN. Nanotwins and stacking faults (SFs) are marked. Inset, a selected area electron diffraction pattern. **e**, HRTEM image of nt-cBN showing Shockley partial dislocations (red triangles) emitted from grain boundaries (GB) and a high density of stacking faults in twin domains. TB, twin boundary. **f**, Enlarged HRTEM image of **e**, showing the orientation relationship between adjacent nanograins.

density of dislocations at grain boundaries, as well as a high density of stacking faults within the twin domains. Narrow and sharp high-angle grain boundaries about five or six atomic layers thick were frequently present (Fig. 1e, f). All nanograins contained densely spaced lamellar {111} twins (Fig. 2a, b). Figure 2c depicts the twin-thickness distribution derived from 726 counts of 60 grains based on TEM and HRTEM images. Twin thicknesses were predominantly below 10 nm, with an average of 3.8 nm. As a result of the low excess energy of twin boundaries, variations in duration (from several minutes to half an hour) during the HPHT treatments did not noticeably affect the average and distribution of twin thickness, indicating that the nanotwins are highly stable.

The H_V of bulk nt-cBN samples was measured with a standard square-pyramidal diamond indenter. Reliable hardness values are best determined from the asymptotic hardness region through a well-controlled indentation process[19,20]. We recorded the variations in H_V with respect to a series of applied loads (Fig. 3). The asymptotic hardness obtained at loads above 3 N reached an extremely high value of 108 GPa, which is the highest hardness reported so far for polycrystalline cBN and exceeds even that of synthetic diamond. Overall, the hardness measured for different nt-cBN samples synthesized between 1,800 and 1,950 °C at 12–15 GPa ranged from 95 GPa to 108 GPa. For direct comparison, we measured the hardness of commercial cBN single crystals with the same technique and obtained values lower than 49 GPa with loads of up to 4.9 N, above which the crystals fractured. The results are in good agreement with other reported hardness data (30–43 GPa) on cBN single crystals[21]. Our measured Knoop and nanoindentation hardnesses of nt-cBN samples are 77.7 and 98.5 GPa (Supplementary Table 1), respectively.

Radial cracks formed in bulk nt-cBN at loads of 9.8 and 19.6 N were used to determine the fracture toughness (K_{IC}) with the equation[22] $K_{IC} = 7.42 \times 10^{-2} F/L^{1.5}$, where L and F are the length of the crack and

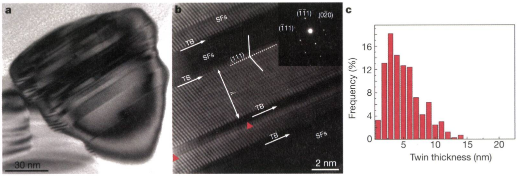


Figure 2 | Microstructure of synthetic nt-cBN. **a**, Bright-field TEM image of a nanograin. **b**, HRTEM image and corresponding selected area electron diffraction pattern (inset) along the [101] zone axis of the nanograin shown in **a**. Lamellar nanotwins with various thicknesses (λ) are present in the nanograin. The twinning plane is of the {111} type, and the lattice fringe angles across the twin plane are 70.53°. Stacking faults and Shockley dislocations (red triangles) are labelled. **c**, Thickness distribution of the twins measured from TEM and HRTEM images with an average twin thickness of 3.8 nm.

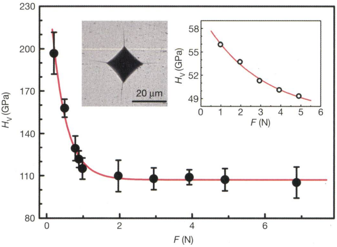


Figure 3 | The H_V of an nt-cBN bulk sample as a function of applied load (*F*). The H_V of the nt-cBN bulk decreases from ~196 GPa at 0.2 N to its asymptotic value, 108 GPa, beyond 3 N. Error bars indicate s.d. ($n = 5$). Left inset, an optical micrograph of the Vickers indentation with cracks produced at a load of 19.6 N. Right inset, the H_V–F curve of a 0.3-mm cBN single crystal. H_V does not reach its asymptotic value, and the crystal fractures when F exceeds 4.9 N.

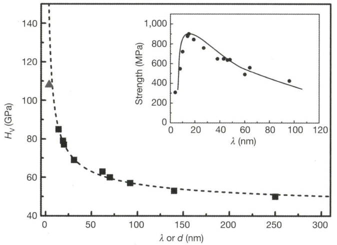


Figure 4 | H_V as a function of average grain size (*d*) or twin thickness (*λ*) for polycrystalline cBN bulk materials. Experimental data for nt-cBN bulk material (triangle) and for ng-cBN bulk materials[1] (squares) are shown. Using equation (1) to fit the experimental data, we obtained $H_V = 42.6 + 126d^{-1/2} + 130.7d^{-1}$. k (126 GPa nm$^{1/2}$) is taken from a previously reported value[1]. The fitted C (130.7 ± 16.8 GPa nm) parameter characterizing the quantum confinement effect is in excellent agreement with the theoretical value (136 GPa nm) from $211N_e^{1/3}\exp(-1.191f_i)$ for cBN. This coincidence is not accidental and may provide proof of the existence of the quantum confinement effect in the synthesized nt-cBN. Inset, the yield strength as a function of λ for nt-Cu, in which the critical λ is about 15 nm (ref. 15).

the applied load, respectively. The resulting K_{IC} reached a remarkably high value of 12.7 MPa m$^{1/2}$, which is about 4.5-fold that of cBN single crystal (2.8 MPa m$^{1/2}$) and almost twice that of micrometre-grain-sized cBN (6.8 MPa m$^{1/2}$) (ref. 2). This K_{IC} is 21% greater than that of ng-cBN (10.5 MPa m$^{1/2}$) (ref. 2) and 27% greater than the fracture toughness (10 MPa m$^{1/2}$) of commercial cemented tungsten carbide, a material widely used in the industry as a cutting tool[23].

The thermal stability of nt-cBN was characterized by using heat flow curves measured in air by differential scanning calorimetry (Supplementary Fig. 3). An onset oxidation temperature of ~1,294 °C was determined from the exothermic trough, which is higher than those of single-crystal cBN (~1,103 °C), ng-cBN (~1,187 °C) (ref. 2) and commercial polycrystalline cBN (~1,000 °C) (ref. 4). The high thermal stability of nt-cBN samples, apparently due to the nanotwinned structure, allows higher-temperature industrial applications.

Figure 4 shows that the hardness of polycrystalline cBN increases monotonically with decreasing microstructural size (grain size for grains and/or twin thickness for twins). On the basis of our recent theoretical model[20], the hardness dependence of the grain size (d) for cBN obeys

$$H_V = H_{HP} + H_{QC} = H_0 + kd^{-1/2} + Cd^{-1} \qquad (1)$$

where $H_{HP}\,(= H_0 + kd^{-1/2})$ and $H_{QC}\,(= Cd^{-1})$ represent dislocation-related hardening based on the Hall–Petch effect[9,10] and bandgap-related hardening based on the speculated quantum confinement effect[24], respectively. H_0 is the single-crystal hardness and k is a material constant. C is a material-specific parameter equal to zero for metals, and equal to $211N_e^{1/3}\exp(-1.191f_i)$ for covalent materials[20], where N_e is the valence electron density and f_i is the Phillips ionicity of the chemical bond[25]. Both Hall–Petch and quantum confinement effects have been used to account for the hardening behaviour in ng-cBN with grain sizes down to 14 nm (ref. 1).

Our present results show that these two hardening mechanisms remain valid when the twin thickness is significantly reduced to 3.8 nm for cBN (Fig. 4). This finding on covalent cBN is in sharp contrast with those in metals, for which both yield strength and hardness decrease significantly (inset of Fig. 4) and the Hall–Petch hardening mechanism becomes invalid[14,15,26] when the twin thickness or grain size decreases to a critical value (typically ~10–15 nm). The continuously hardening behaviour with decreasing microstructural sizes (down to 3.8 nm) in cBN may be rationalized as follows. For nanotwins with thicknesses below 3.8 nm, the quantum confinement effect (which is inoperative in metals) becomes dominantly large according to equation (1) (see also the caption of Fig. 4). At a twin thickness of 3.8 nm, each twin lamella contains only about ten unit cells of cBN. This featured nanoscale geometry, as well as the very strong covalent B–N bonding, severely confines the migration of twin boundaries, which is known to induce hardness softening in nanotwinned metals[26].

The use of oBN nanoparticles as precursors is the key to synthesizing nt-cBN, and it highlights the important structural effects of spherical atomic arrangement in oBN on the formation of nanotwinned substructures. It has been established that coarse-grained graphite-like BN precursors can transform into grained cBN or wurtzite BN through martensitic transformation by means of puckering or buckling of the planar BN layers[3]. In our synthesis, the oBN precursors that we used contain naturally puckered BN layers (Fig. 1b), which make the transformation into cBN easier. The high concentration of puckered layers and defects (here, stacking faults; Fig. 1a) in oBN precursors provides ideal sites for the nucleation of cBN. Moreover, the specific orientation relations between coherent nucleus cBN and parent phase oBN must be reserved during the reconstructive phase transition[27]. As a result, multiple parallel laminated nanotwins form at the sub-grain level. High-density domain boundaries with ideal boundary structures in nanotwins make nt-cBN the most desirable superhard material yet achieved. This new material possesses the combined advantages of ultrahigh hardness, high thermal stability, and extremely high fracture toughness, providing a wide range of new industrial applications.

Our findings introduce a new strategy and direction in the quest for superhard materials. The synthesis of nanotwinned substructures by selecting suitable starting materials and optimal P–T conditions, and the use of hardening effects from interlocked nanotwins, are essential. On the basis of our results, if nanotwins at similar scales can be reproduced in polycrystalline diamond (isostructural to cBN and known for twinning in large crystals), an ultrahigh hardness exceeding that of ng-diamond would be achievable.

METHODS SUMMARY

Pure oBN nanospheres with diameters about 30–150 nm were synthesized with a chemical vapour deposition technique[28]. A schematic icosahedral model of a five-shell oBN nanoparticle was relaxed from the nested bucky onion of $B_{30}N_{30}$, $B_{120}N_{120}$, $B_{270}N_{270}$, $B_{480}N_{480}$ and $B_{750}N_{750}$ fullerenes. Geometric optimization was performed with the Forcite module of Materials Studio software, and the Universal Force Field was used. HPHT experiments were performed with a 10-MN two-stage large-volume multi-anvil system at Yanshan University, identical to that described in ref. 29. The standard COMPRES 10/5 sample assembly, consisting of a 10-mm spinel ($MgAl_2O_4$) octahedron with a Re heater and a $LaCrO_3$ thermal insulator, was used[30]. Temperature was measured *in situ* with Type C W-Re thermocouples, and pressures were estimated from previously obtained calibration curves at different temperatures[30]. Recovered bulk samples were about 2 mm in diameter and 0.2–0.5 mm in height. Microstructures of the bulks were characterized by TEM (JEOL-2010F) with an accelerating voltage of 200 kV at the University of New Mexico and TEM (JEM-2010) with an accelerating voltage of 200 kV at Yanshan University. Phase identification was performed with an X-ray powder diffractometer (D8 Discover) with Cu K_α radiation, and by Raman spectroscopy (Renishaw inVia). H_V and Knoop hardness (H_K) were measured with a microhardness tester (KB 5 BVZ). At least five hardness data points were taken for each load, and the hardness values were determined from the asymptotic-hardness region. Nanohardness (H_N) and Young's modulus (E) were measured with a Hysitron TI 950 TriboIndentor with Berkovich indenter at a peak load of 50 mN (see Supplementary Fig. 4 for details). Studies on oxidation resistance were performed in air by differential scanning calorimetry (NETZSCH STA 449 C) with a heating rate of 10 °C min^{-1} in the temperature range 20–1,500 °C.

Received 20 May; accepted 29 October 2012.

1. Dubrovinskaia, N. *et al.* Superhard nanocomposite of dense polymorphs of boron nitride: noncarbon material has reached diamond hardness. *Appl. Phys. Lett.* **90,** 101912 (2007).
2. Solozhenko, V. L., Kurakevych, O. O. & Le Godec, Y. Creation of nanostuctures by extreme conditions: high-pressure synthesis of ultrahard nanocrystalline cubic boron nitride. *Adv. Mater.* **24,** 1540–1544 (2012).
3. Britun, V. F. & Kurdyumov, A. V. Mechanisms of martensitic transformations in boron nitride and conditions of their development. *High Press. Res.* **17,** 101–111 (2000).
4. Sumiya, H., Uesaka, S. & Satoh, S. Mechanical properties of high purity polycrystalline cBN synthesized by direct conversion sintering method. *J. Mater. Sci.* **35,** 1181–1186 (2000).
5. Dub, S. N. & Petrusha, I. A. Mechanical properties of polycrystalline cBN obtained from pyrolytic gBN by direct transformation technique. *High Press. Res.* **26,** 71–77 (2006).
6. Wentorf, R. H., DeVries, R. C. & Bundy, F. P. Sintered superhard materials. *Science* **208,** 873–880 (1980).
7. Liew, W. Y. H., Yuan, S. & Ngoi, B. K. A. Evaluation of machining performance of STAVAX with PCBN tools. *Int. J. Adv. Manuf. Technol.* **23,** 11–19 (2004).
8. Krauss, A. R. *et al.* Ultrananocrystalline diamond thin films for MEMS and moving mechanical assembly devices. *Diamond Related Materials* **10,** 1952–1961 (2001).
9. Hall, E. O. The deformation and ageing of mild steel: III. Discussion of results. *Proc. Phys. Soc. B* **64,** 747–753 (1951).
10. Petch, N. J. The cleavage strength of polycrystals. *J. Iron Steel Inst.* **174,** 25–28 (1953).
11. Irifune, T., Kurio, A., Sakamoto, S., Inoue, T. & Sumiya, H. Ultrahard polycrystalline diamond from graphite. *Nature* **421,** 599–600 (2003).
12. Sumiya, H. & Irifune, T. Hardness and deformation microstructures of nano-polycrystalline diamonds synthesized from various carbons under high pressure and high temperature. *J. Mater. Res.* **22,** 2345–2351 (2007).
13. Schiøtz, J., Di Tolla, F. D. & Jacobsen, K. W. Softening of nanocrystalline metals at very small grain sizes. *Nature* **391,** 561–563 (1998).
14. Schiøtz, J. & Jacobsen, K. W. A maximum in the strength of nanocrystalline copper. *Science* **301,** 1357–1359 (2003).
15. Lu, L., Chen, X., Huang, X. & Lu, K. Revealing the maximum strength in nanotwinned copper. *Science* **323,** 607–610 (2009).
16. Pande, C. S. & Cooper, K. P. Nanomechanics of Hall–Petch relationship in nanocrystalline materials. *Prog. Mater. Sci.* **54,** 689–706 (2009).
17. He, L. L., Akaishi, M. & Horiuchi, S. Structural evolution in boron nitrides during the hexagonal–cubic phase transition under high pressure at high temperature. *Microsc. Res. Tech.* **40,** 243–250 (1998).
18. Oku, T., Hiraga, K., Matsuda, T., Hirai, T. & Hirabayashi, M. Twin structures of rhombohedral and cubic boron nitride prepared by chemical vapor deposition method. *Diamond Related Materials* **12,** 1138–1145 (2003).
19. Brazhkin, V. *et al.* What does 'harder than diamond' mean? *Nature Mater.* **3,** 576–577 (2004).
20. Tian, Y., Xu, B. & Zhao, Z. Microscopic theory of hardness and design of novel superhard crystals. *Int. J. Refract. Met. Hard Mater.* **33,** 93–106 (2012).
21. Harris, T. K., Brookes, E. J. & Taylor, C. J. The effect of temperature on the hardness of polycrystalline cubic boron nitride cutting tool materials. *Int. J. Refract. Met. Hard Mater.* **22,** 105–110 (2004).
22. Solozhenko, V. L., Dub, S. N. & Novikov, N. V. Mechanical properties of cubic BC_2N, a new superhard phase. *Diamond Related Materials* **10,** 2228–2231 (2001).
23. Berger, C., Scheerer, H. & Ellermeier, J. Modern materials for forming and cutting tools—overview. *Materialwiss. Werkstofftech.* **41,** 5–17 (2010).
24. Tse, J. S., Klug, D. D. & Gao, F. M. Hardness of nanocrystalline diamonds. *Phys. Rev. B* **73,** 140102 (2006).
25. Gao, F. M. *et al.* Hardness of covalent crystals. *Phys. Rev. Lett.* **91,** 015502 (2003).
26. Li, X., Wei, Y., Lu, L., Lu, K. & Gao, H. Dislocation nucleation governed softening and maximum strength in nano-twinned metals. *Nature* **464,** 877–880 (2010).
27. Oleinik, G. S., Petrusha, I. A., Danilenko, N. V., Kotko, A. V. & Shevchenko, S. A. Crystal-oriented mechanism of dynamic recrystallization nucleation in cubic boron nitride. *Diamond Related Materials* **7,** 1684–1692 (1998).
28. Tang, C. C., Bando, Y., Huang, Y., Zhi, C. Y. & Golberg, D. Synthetic routes and formation mechanisms of spherical boron nitride nanoparticles. *Adv. Funct. Mater.* **18,** 3653–3661 (2008).
29. Wang, Y. *et al.* The large-volume high-pressure facility at GSECARS: a 'Swiss-army-knife' approach to synchrotron-based experimental studies. *Phys. Earth Planet. Inter.* **174,** 270–281 (2009).
30. Leinenweber, K. D. *et al.* Cell assemblies for reproducible multi-anvil experiments (the COMPRES assemblies). *Am. Mineral.* **97,** 353–368 (2012).

Supplementary Information is available in the online version of the paper.

Acknowledgements We thank J. K. Yu for help with the differential scanning calorimetry measurements. Y.J.T. and Z.Y.L. acknowledge financial support from the Ministry of Science and Technology of China (grants 2011CB808205 and 2010CB731605), Y.J.T., D.L.Y., Y.M.M. and J.L.H. are grateful for financial support from the National Natural Science Foundation of China (grants 51121061, 51172197, 11025418 and 91022029), and Y.B.W. acknowledges financial support from the US National Science Foundation (EAR-0968456).

Author Contributions Y.J.T. conceived the project. Y.J.T., B.X., D.L.Y. and Y.B.W. designed the experiments. C.C.T. synthesized oBN precursors, Y.J.T., B.X., D.L.Y., Y.B.W., Y.F.G., K.L. and Z.S.Z. performed the HPHT experiments, Y.B.J. and W.T.H. performed TEM observations, and B.W. performed molecular dynamics simulations. Y.J.T., B.X., D.L.Y., Y.M.M., Y.B.W., L.-M.W., J.L.H. and Z.Y.L. analysed the data. Y.J.T., B.X., Y.M.M. and Y.B.W. co-wrote the paper. Y.J.T., B.X. and D.L.Y. contributed equally to the study. All authors discussed the results and commented on the manuscript.

Author Information Reprints and permissions information is available at www.nature.com/reprints. The authors declare no competing financial interests. Readers are welcome to comment on the online version of the paper. Correspondence and requests for materials should be addressed to Y.J.T. (fhcl@ysu.edu.cn).

doi:10.1038/nature11785

Multidecadal variability in East African hydroclimate controlled by the Indian Ocean

Jessica E. Tierney[1,2], Jason E. Smerdon[2], Kevin J. Anchukaitis[1,2] & Richard Seager[2]

The recent decades-long decline in East African rainfall[1] suggests that multidecadal variability is an important component of the climate of this vulnerable region. Prior work based on analysing the instrumental record implicates both Indian[2] and Pacific[1] ocean sea surface temperatures (SSTs) as possible drivers of East African multidecadal climate variability, but the short length of the instrumental record precludes a full elucidation of the underlying physical mechanisms. Here we show that on timescales beyond the decadal, the Indian Ocean drives East African rainfall variability by altering the local Walker circulation, whereas the influence of the Pacific Ocean is minimal. Our results, based on proxy indicators of relative moisture balance for the past millennium paired with long control simulations from coupled climate models, reveal that moist conditions in coastal East Africa are associated with cool SSTs (and related descending circulation) in the eastern Indian Ocean and ascending circulation over East Africa. The most prominent event identified in the proxy record—a coastal pluvial from 1680 to 1765—occurred when Indo-Pacific warm pool SSTs reached their minimum values of the past millennium. Taken together, the proxy and model evidence suggests that Indian Ocean SSTs are the primary influence on East African rainfall over multidecadal and perhaps longer timescales.

The 2010–2011 drought in the Horn of Africa, by some measures the worst drought in 60 years[3], is a reminder that rainfall in this politically and socioeconomically vulnerable region can fluctuate dramatically. Prevailing La Niña conditions in the tropical Pacific were partly to blame; East African rainfall is teleconnected to the El Niño/Southern Oscillation[4,5] (ENSO), and the Horn of Africa experiences droughts during La Niña events and pluvials during El Niño events. However, it is debated whether the failure of the 'long rains' (the rainy season of March, April and May) in 2011—which exacerbated the drought—is related to decadal variability in the Indo-Pacific region[1] or anthropogenic forcing[1,2,6]. It is critical for us to understand the character and mechanisms that drive decadal to centennial shifts in East African rainfall if we are to evaluate future regional projections of drought frequency and food security, but the short length of the instrumental record fundamentally limits our ability to understand variability on these timescales using observational data alone.

Palaeoclimate records in East Africa from the past millennium have the potential to extend the instrumental record and reveal mechanisms driving low-frequency climate variability. At present, annually resolved, absolutely dated terrestrial archives such as tree rings are sparse or underdeveloped in the region, but numerous lake basins in East Africa provide sedimentary archives with sufficient accumulation rates to resolve past-millennium climate[7]. Lake archives nevertheless have a fundamental limitation: they are not absolutely dated and, if constrained by radiocarbon (^{14}C) dating, can have a relatively large (~50–100-yr) temporal uncertainty due to compounded analytical and calibration errors. This uncertainty can make the identification of shared trends between different site archives challenging, especially within the time frame of the past millennium.

Here we synthesize lacustrine hydroclimatic proxy records from East Africa using a Monte Carlo empirical orthogonal function (MCEOF) approach[8] to develop a spatiotemporal view of regional water balance during the past millennium that accounts for time uncertainty. We use proxy data from seven different lake basins (Fig. 1a; see Methods Summary for proxy selection criteria) that include charcoal, run-off indicators, lake level reconstructions and leaf wax hydrogen isotopes. The leading MCEOF of the proxy data explains $32 \pm 6\%$ (2σ) of the variance in the data and describes the primary regional mode of hydroclimatic variance during the past ~700 years (Fig. 1b). For sites that load positively on MCEOF1, the mode describes a pattern of late-medieval drought (1300 to 1400) followed by a gradual transition towards wetter conditions, culminating in peak pluvial conditions from about 1700 to 1750, and a subsequent abrupt transition back towards drier conditions that persist until modern times (1950; Fig. 1b). Sites that load negatively on MCEOF1 show the opposite pattern and are dry during the eighteenth-century pluvial.

Some of the major features of MCEOF1, such as the medieval drought and the eighteenth-century pluvial, have been discussed and identified in previous site-specific studies[9,10] and are also evident in some of the individual proxy data time series (Fig. 1a). Our analysis, however, identifies the spatial patterns of these features as well as their associated uncertainties. The spatial loading patterns of MCEOF1 clearly indicate that the sites closer to the eastern coastal/Horn of Africa region vary in antiphase with the interior Rift Valley sites (Fig. 1c). Although EOF analysis imposes a constraint of orthogonality that can complicate the interpretation of loading patterns in a climatic sense[11], this Horn–Rift dichotomy is consistent with convergence anomalies that occur in East Africa in response to altered Indian Ocean SST gradients[12], and also approximates the spatial pattern of the first EOF of 10-yr low-pass-filtered instrumental rainfall data in the region (Fig. 1 and Supplementary Information). These features suggest that the loading pattern in MCEOF1 represents a real and important aspect of climate variability in East Africa that has dominated hydroclimate during the past millennium on the multidecadal timescale.

To explore the possible climatic mechanisms driving this variability, we analyse the relationship between annual precipitation in easternmost Africa—the coastal area that loads positively on MCEOF1—and SSTs in millennium-long control simulations conducted with fully coupled, atmosphere–ocean general circulation models (AOGCMs). These simulations provide sufficient degrees of freedom to investigate unforced, decadal to centennial climate variability, something that cannot be achieved with instrumental data or model simulations spanning only a few centuries. We use the 1,300- and 3,000-yr control runs from the US National Center for Atmospheric Research (NCAR) CCSM4[13] AOGCM and the Geophysical Fluid Dynamics Laboratory (GFDL) CM2.1[14,15] AOGCM, respectively. Both models correctly simulate the ENSO-driven teleconnection between tropical Indo-Pacific SSTs and East African rainfall apparent in the instrumental record, although the correlation is stronger in the former than in the latter (Supplementary Fig. 3). To isolate low-frequency relationships, we apply a 50-yr low-pass filter to both the SST and the precipitation field, in line with the highest frequency recovered in MCEOF1 (Supplementary Fig. 4), and calculate the field correlations (Fig. 2). We find that although the central and eastern Pacific influence on East African rainfall is strong in the annual fields, it disappears on

---

[1]Woods Hole Oceanographic Institution, Woods Hole, Massachusetts 02543, USA. [2]Lamont-Doherty Earth Observatory of Columbia University, Palisades, New York 10964, USA.

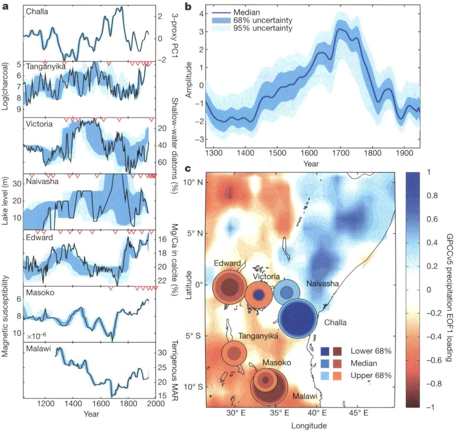


Figure 1 | Results of the MCEOF proxy synthesis. **a**, The seven proxy data time series used[9,10,20–28] (black lines indicate data on their published age models) and their associated time uncertainties (blue shading: light, 68%; dark, 95%). Red triangles denote chronological constraints (except for the varved Malawi and Challa data). Lake Challa's '3-proxy PC1' is the first principal component of three hydroclimate proxies (Supplementary Information). Vertical axes are oriented such that wet conditions plot upwards. MAR, mass accumulation rate. **b**, MCEOF1, shown with the median (solid blue line) and 68% and 95% two-tailed uncertainty bounds (blue shadings) empirically determined by 10,000 simulations. **c**, Proxy loadings on MCEOF1 (coloured circles), superimposed on the loading pattern (background colours) of the first EOF of the 10-yr low-pass-filtered instrumental precipitation field (GPCCv5[29]). Circle diameter represents the loading value; circle colour represents the lower 68% bound (inner circle), median (middle circle) and upper 68% bound (outer circle).

decadal and longer timescales (Fig. 2 and Supplementary Fig. 5). The only significant correlations that remain in the 50-yr low-pass-filtered field represent the association of wet conditions with warm SSTs in the western Indian Ocean and cool SSTs in the eastern Indian Ocean and western Pacific warm pool (Fig. 2).

Changes in model atmospheric vertical velocity are consistent with the changes in SSTs and provide insight into the mechanisms linking the surface ocean to rainfall variability. Figure 3 shows longitude–height cross-sections of vertical velocity correlated with East African precipitation for the GFDL CM2.1 model. For the unfiltered data (Fig. 3a), wet conditions in East Africa are associated with a reorganization of the Walker circulation throughout the tropics, with anomalous rising motion over the central and eastern equatorial Pacific, descending motion over the far western Pacific and eastern Indian Ocean, and another cell of anomalous rising motion over the western Indian Ocean and East Africa. These correlations are typical of ENSO and are consistent with patterns of low-level wind anomalies driven by the SST gradients and associated atmospheric heating anomalies. In contrast, the low-pass-filtered vertical velocity correlations show that on multidecadal and longer timescales the tropical Pacific influence is considerably weaker (Fig. 3b). Instead, East African precipitation anomalies are controlled by a Walker circulation anomaly that is localized over the Indian Ocean, with wet conditions associated with ascending motion over East Africa and the western Indian Ocean and descending motion over the eastern Indian Ocean.

Collectively, these simulations imply that changes in Indian Ocean SSTs or the Indian Ocean SST gradient—and not the Pacific SST gradient—are the dominant influence on East African rainfall on multidecadal and longer timescales. Such changes in SSTs act to weaken or strengthen the Walker circulation in the Indian Ocean basin, thereby causing multidecadal pluvials or droughts, respectively. The Pacific influence on East African rainfall decays beyond the 10-yr timescale (Supplementary Fig. 5), probably reflecting the waning influence of interannual ENSO. Indeed, analysis of the power spectra of the simulated SSTs in the Pacific and Indian ocean basins, respectively, suggests that the central Pacific Ocean has relatively little spectral power in the multidecadal band whereas both the eastern and western Indian Ocean basins retain relatively more low-frequency power (Supplementary Fig. 6). These low-frequency oscillations are clearly capable of affecting the Walker circulation in the Indian Ocean basin independently of the Pacific, lending credence to, as well as extending, previous work suggesting that the Indian Ocean directly affects East African hydroclimate even in the interannual band[12,16].

Changing Indian Ocean SST gradients may have dictated the evolution of East African hydroclimate during the past millennium. Independent proxy evidence of Indian Ocean SSTs would provide the most salient means of investigating this, but proxy SST records from the Indian Ocean basin over the past millennium are so far limited to coral archives that span at most the past few centuries. However, a recent sedimentary reconstruction from the Makassar strait provides a continuous SST record for the western Pacific warm pool over the past millennium[17] and, given the known influence of the Indonesian Throughflow on southeastern Indian Ocean SSTs[16], provides an indirect estimate of temperature variability in the easternmost portion of the Indian Ocean basin. This sedimentary reconstruction illustrates a remarkable similarity between the temporal evolution of Makassar SSTs and East African climate, with the coldest SSTs of the past millennium coeval with the coastal pluvial evident in MCEOF1 (Fig. 4). These palaeoclimate data support the mechanistic link between eastern Indian Ocean SSTs and East African rainfall seen in the AOGCM control runs, wherein cool SSTs in the eastern Indian Ocean are associated with wet conditions along the East African coast.

Determination of the role of western Indian Ocean SSTs, which are thought to be critical in terms of promoting local convergence anomalies on the interannual basis[12], awaits long and continuous proxy SST data from the western Indian Ocean. However, it is dynamically consistent with our AOGCM results that cooler conditions in the western Pacific warm pool during the late seventeenth century and

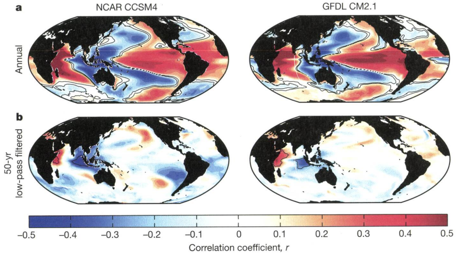


Figure 2 | Field correlations between simulated East African precipitation and SSTs. East African precipitation was averaged over the area spanned by 5° S–7° N and 36°–46° E, and then correlated with SSTs from long control simulations conducted with the NCAR CCSM4 AOGCM and the GFDL CM2.1 AOGCM, respectively. **a**, Annual correlations; **b**, 50-yr low-pass-filtered correlations. Black contour lines demarcate regions in which a null hypothesis of no correlation can be rejected at the 5% level. *r*, Pearson correlation coefficient.

the early eighteenth century helped reduced the east–west gradient in Indian Ocean SSTs, weakened the Walker circulation and promoted pluvial conditions in coastal East Africa and drought in the interior Rift Valley. As the AOGCM simulations demonstrate, such multidecadal modulations of SSTs and East African climate can occur as part of natural, unforced climate variability, leaving open the possibility that the pluvial event in coastal Africa was at least in part a result of internal variability. However, the pluvial event and the cool conditions in the warm pool do occur during one of the coldest intervals of the Northern Hemisphere Little Ice Age[18] and thus may be part of a global climate reorganization in response to radiative forcing.

This work presents an analysis of the spatial and temporal character of East African hydroclimate on decadal to centennial timescales, based on both palaeoclimate data and climate model simulations. In contrast to the dominant impact of ENSO on interannual rainfall variability, variations in the SST gradient across the Indian Ocean—which alter the Walker circulation—seem to be the principal control on hydroclimate in East Africa on multidecadal timescales. Whereas previous interpretations of palaeoclimate data from East Africa typically invoked ENSO or ENSO-like controls on East African precipitation[10], our analysis suggests that palaeoclimate data that preferentially or exclusively record multidecadal or longer climate variability may be better understood in the context of the Indian Ocean, which has its own agency at lower frequencies.

The modelling simulations demonstrate that multidecadal oscillations in Indian Ocean SSTs can arise in response to internal climate

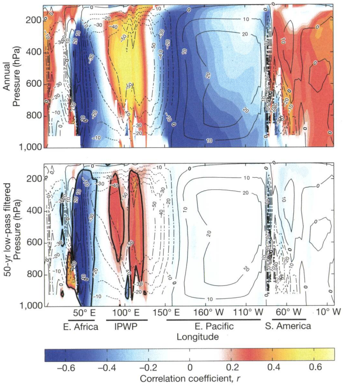


Figure 3 | Field correlations between simulated East African precipitation and vertical velocity. 1Data are from the GFDL CM2.1 simulation. Precipitation was averaged as in Fig. 2 and then correlated with vertical velocity (ω; positive values indicate descending motion; averaged over the 5° S–5° N latitude band) across the atmospheric pressure field. **a**, Annual field, showing only values for which a null hypothesis of no correlation can be rejected at the 5% level. **b**, 50-yr low-pass-filtered field. Heavy black contours demarcate regions in which a null hypothesis of no correlation can be rejected at the 5% level. Light contours represent mean-annual simulated ω (hectopascals per day), indicating the locations of the ascending and descending branches of the Walker circulation. IPWP, Indo-Pacific warm pool.

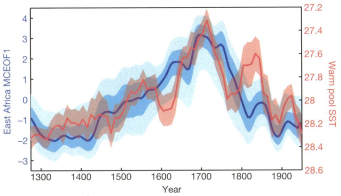


Figure 4 | Comparison between East Africa MCEOF1 (blue) and an SST reconstruction from the Makassar strait in the western Pacific warm pool (orange)[17]. The right-hand axis is flipped such that cooler conditions plot upwards. Shading on MCEOF1 indicates the 68% (dark) and 95% (light) two-tailed uncertainty bounds, as in Fig. 1. Shading on Makassar SSTs indicates the standard error on the binned proxy SST data as described in ref. 17.

variability, but the palaeoclimate data hint at a forced response in that a cooler western Pacific warm pool and a coastal East African pluvial are associated with the Little Ice Age. Efforts to explain the recent decadal drought in the region and project future hydroclimatic change, whether forced or unforced, must acknowledge the existence of potent low-frequency hydroclimate variability related to the Indian Ocean that is not detectable from the instrumental record alone. Present climate models predict that East Africa will get wetter as a consequence of increasing concentrations of greenhouse gases[19], but the region has in fact gotten drier in recent decades[1]. This discrepancy could be explained either if a naturally occurring dry period, possibly of the type identified here, is obscuring a radiatively forced wetting trend or if the model projections are incorrect. More reliable projections of hydroclimate in this vulnerable region will require a better understanding of regional hydroclimate variability on decadal and longer timescales, how the regional climate responds to radiative forcings and how these two factors will combine to determine the future hydroclimate of East Africa.

METHODS SUMMARY

Seven palaeoclimate records of relative water balance from the East African region were included in the MCEOF synthesis, on the basis of the following criteria: the proxy primarily reflects hydroclimate; the proxy was measured at a time interval of 50 years or less; the age model is based on at least seven age–depth tie-points; there is at least one proxy data point representative of modern (post-1950) conditions; and the archive has a well-constrained stratigraphy free of reworking, large turbidites and sedimentary hiatuses. The MCEOF method, which is fully described in ref. 8, was applied to the data with the number of simulations set to 10,000 (ref. 8). Palaeoclimate data and age constraints were treated as described in ref. 8 with the exception of data from Lake Challa, in which case the first principal component of the three available hydroclimate proxy time series was used (Supplementary Information).

The GFDL CM2.1 model's[14] pre-industrial control simulation spans 3,000 years and uses an atmospheric resolution of 2° latitude by 2° longitude by 24 vertical levels and an oceanic resolution of 1° latitude (increasing to 1/3° near the equator) by 1° longitude by 50 vertical levels. The NCAR CCSM4 model's[13] pre-industrial control simulation spans 1,300 years and uses an atmospheric resolution of 1° latitude by 1° longitude by 26 vertical layers and an oceanic resolution of 0.54° latitude (increasing to 0.27° near the equator) by 1.11° longitude by 60 vertical levels. For model analyses, we averaged precipitation over the area spanned by 5° S–7° N and 36°–46° E. Annual precipitation, SST and vertical velocity (for the GFDL simulation) fields were filtered with an eight-point Butterworth filter before calculating the correlation coefficients of the low-pass-filtered fields in Figs 2 and 3. All annual fields were calculated from monthly means. For further details regarding the model analyses, see Supplementary Information.

Data from this paper is available from NOAA's World Data Center for Paleoclimatology (www.ncdc.noaa.gov/paleo).

Received 30 July; accepted 8 November 2012.

1. Lyon, B. & DeWitt, D. A recent and abrupt decline in the East African long rains. *Geophys. Res. Lett.* **39,** L02702 (2012).
2. Williams, A. & Funk, C. A westward extension of the warm pool leads to a westward extension of the Walker circulation, drying eastern Africa. *Clim. Dyn.* **37,** 2417–2435 (2011).
3. Famine Early Warning System Network. East Africa: past year one of the driest on record in the eastern Horn. http://www.fews.net/docs/Publications/FEWS NET EA_Historical drought context_061411pdf (2011).
4. Nicholson, S. & Entekhabi, D. The quasi-periodic behavior of rainfall variability in Africa and its relationship to the Southern Oscillation. *Meteorol. Atmos. Phys.* **34,** 311–348 (1986).
5. Ropelewski, C. & Halpert, M. Global and regional scale precipitation patterns associated with the El Niño/Southern Oscillation. *Mon. Weath. Rev.* **115,** 1606–1626 (1987).
6. Funk, C. *et al.* Warming of the Indian Ocean threatens eastern and southern African food security but could be mitigated by agricultural development. *Proc. Natl Acad. Sci. USA* **105,** 11081–11086 (2008).
7. Verschuren, D. in *Past Climate Variability Through Europe and Africa* (eds Battarbee, R. W., Gasse, F. & Stickley, C. E.) 219–256 (Springer, 2004).
8. Anchukaitis, K. J. & Tierney, J. E. Identifying coherent spatiotemporal modes in time-uncertain proxy paleoclimate records. *Clim. Dyn.* advance online publication, http://dx.doi.org/10.1007/s00382-012-1483-0 (26 August 2012).
9. Verschuren, D., Laird, K. R. & Cumming, B. F. Rainfall and drought in equatorial East Africa during the past 1,100 years. *Nature* **403,** 410–414 (2000).
10. Russell, J. M. & Johnson, T. C. Little Ice Age drought in equatorial Africa: Intertropical Convergence Zone migrations and El Niño-Southern Oscillation variability. *Geology* **35,** 21–24 (2007).
11. Dommenget, D. & Latif, M. A cautionary note on the interpretation of EOFs. *J. Clim.* **15,** 216–225 (2002).
12. Ummenhofer, C. C., Gupta, A. S., England, M. H. & Reason, C. J. C. Contributions of Indian Ocean sea surface temperatures to enhanced East African rainfall. *J. Clim.* **22,** 993–1013 (2009).
13. Gent, P. *et al.* The community climate system model version 4. *J. Clim.* **24,** 4973–4991 (2011).
14. Delworth, T. *et al.* GFDL's CM2 global coupled climate models. Part I: formulation and simulation characteristics. *J. Clim.* **19,** 643–674 (2006).
15. Wittenberg, A. Are historical records sufficient to constrain ENSO simulations? *Geophys. Res. Lett.* **36,** L12702 (2009).
16. Black, E., Slingo, J. & Sperber, K. R. An observational study of the relationship between excessively strong short rains in coastal East Africa and Indian Ocean SST. *Mon. Weath. Rev.* **131,** 74–94 (2003).
17. Oppo, D. W., Rosenthal, Y. & Linsley, B. K. 2,000-year-long temperature and hydrology reconstructions from the Indo-Pacific warm pool. *Nature* **460,** 1113–1116 (2009).
18. Jansen, E. J. *et al.* in *Climate Change 2007: The Physical Science Basis* (eds Solomon, S. *et al.*) 433–498 (Cambridge Univ. Press, 2007).
19. Christensen, J. H. *et al.* in *Climate Change 2007: The Physical Science Basis* (eds Solomon, S. *et al.*) 847–940 (Cambridge Univ. Press, 2007).
20. Verschuren, D. *et al.* Half-precessional dynamics of monsoon rainfall near the East African equator. *Nature* **462,** 637–641 (2009).
21. Tierney, J. E., Russell, J. M., Sinninghe Damsté, J. S., Huang, Y. & Verschuren, D. Late Quaternary behavior of the East African monsoon and the importance of the Congo Air Boundary. *Quat. Sci. Rev.* **30,** 798–807 (2011).
22. Wolff, C. *et al.* Reduced interannual rainfall variability in East Africa during the last ice age. *Science* **333,** 743–747 (2011).
23. Stager, J. C., Ryves, D. B., Cumming, B. F., Meeker, L. D. & Beer, J. Solar variability and the levels of Lake Victoria, East Africa, during the last millenium. *J. Paleolimnol.* **33,** 243–251 (2005).
24. Tierney, J. *et al.* Late-twentieth-century warming in Lake Tanganyika unprecedented since AD 500. *Nature Geosci.* **3,** 422–425 (2010).
25. Garcin, Y. *et al.* Centennial to millennial changes in maar-lake deposition during the last 45,000 years in tropical Southern Africa (Lake Masoko, Tanzania). *Palaeogeogr. Palaeoclimatol. Palaeoecol.* **239,** 334–354 (2006).
26. Garcin, Y. *et al.* Solar and anthropogenic imprints on Lake Masoko (southern Tanzania) during the last 500 years. *J. Paleolimnol.* **37,** 475–490 (2007).
27. Brown, E. T. & Johnson, T. C. Coherence between tropical East African and South American records of the Little Ice Age. *Geochem. Geophys. Geosyst.* **6,** Q12005 (2005).
28. Johnson, T. & McCave, I. Transport mechanism and paleoclimatic significance of terrigenous silt deposited in varved sediments of an African rift lake. *Limnol. Oceanogr.* **53,** 1622–1632 (2008).
29. Rudolf, B., Becker, A., Schneider, U., Meyer-Christoffer, A. & Ziese, M. GPCC status report, December 2011. http://www.dwd.de/bvbw/generator/DWDWWW/Content/Oeffentlichkeit/KU/KU4/KU42/en/Reports_Publications/GPCC_status_report_2010,templateId=raw,property=publicationFile.pdf/GPCC_status_report_2010.pdf (2011).

Supplementary Information is available in the online version of the paper.

Acknowledgements J.E.T. acknowledges the US NOAA Climate and Global Change Postdoctoral Fellowship and NSF OCE-1203892 for support. J.E.S. and R.S. were supported by the NOAA award Global Decadal Hydroclimate Variability and Change (NA10OAR431037). We thank the National Center for Atmospheric Research, who made the CCSM4 control simulation available through the US Earth System Grid (ESG) Center. Support for the ESG is provided by the Office of Science, US Department of Energy, with co-sponsorship from the US NSF. Thanks also to NOAA GFDL for providing the multimillennial control simulation output from the GFDL CM2.1 model, to N. Naik for her on-site support at LDEO and to J. Jungclaus of the Max Planck Institute for providing output from the MPI COSMOS control simulation. This is LDEO contribution number 7641.

Author Contributions J.E.T. designed the study and interpreted the palaeoclimate results. J.E.S. provided the climate modelling results. K.J.A. and J.E.T. designed and implemented the MCEOF method. J.E.S. and R.S. interpreted the dynamical implications of the climate model output. All authors collaborated on the synthesis of the proxy and model data and the writing of the manuscript.

Author Information Reprints and permissions information is available at www.nature.com/reprints. The authors declare no competing financial interests. Readers are welcome to comment on the online version of the paper. Correspondence and requests for materials should be addressed to J.E.T. (tierney@whoi.edu).

LETTER

doi:10.1038/nature11771

Nickel and helium evidence for melt above the core–mantle boundary

Claude Herzberg[1], Paul D. Asimow[2], Dmitri A. Ionov[3,4], Chris Vidito[1], Matthew G. Jackson[5] & Dennis Geist[6]

High ^{3}He/^{4}He ratios in some basalts have generally been interpreted as originating in an incompletely degassed lower-mantle source[1–9]. This helium source may have been isolated at the core–mantle boundary region since Earth's accretion[4–6]. Alternatively, it may have taken part in whole-mantle convection and crust production over the age of the Earth[7–9]; if so, it is now either a primitive refugium at the core–mantle boundary[8] or is distributed throughout the lower mantle[7,9]. Here we constrain the problem using lavas from Baffin Island, West Greenland, the Ontong Java Plateau, Isla Gorgona and Fernandina (Galapagos). Olivine phenocryst compositions show that these lavas originated from a peridotite source that was about 20 per cent higher in nickel content than in the modern mid-ocean-ridge basalt source. Where data are available, these lavas also have high ^{3}He/^{4}He. We propose that a less-degassed nickel-rich source formed by core–mantle interaction during the crystallization of a melt-rich layer or basal magma ocean[5,6], and that this source continues to be sampled by mantle plumes. The spatial distribution of this source may be constrained by nickel partitioning experiments at the pressures of the core–mantle boundary.

Primitive mantle has been estimated to contain 1,960 p.p.m. nickel (Ni)[10], and this is similar to estimates for depleted peridotite[11], which makes up the average mid-ocean-ridge basalt (MORB) source. It is also similar to the Ni content of fertile peridotite (which has less basalt than primitive peridotite but more melt than depleted peridotite so more basalt can be extracted during partial melting) obtained from recent high-quality measurements[12] (Supplementary Fig. 1 and Supplementary section 1), which can be described by:

$$\text{Ni (p.p.m.)} = 68.6\,\text{MgO (weight per cent)} - 630 \qquad (1)$$

Equation (1) fits the Ni contents of 110 fertile and depleted peridotite compositions to 92 p.p.m. ($\pm 2\sigma$); we refer to this as normal Ni peridotite. High-precision olivine analyses from these fertile and depleted peridotites[12] yield 2,800–3,100 p.p.m. Ni and magnesium (Mg)-numbers of 89–92 (that is, 100MgO/(MgO + FeO) in mole per cent; Fig. 1a, Supplementary Table 1 and Supplementary Information section 2). These olivines from mantle peridotite are very similar to primitive olivine phenocryst compositions from East Pacific Rise (EPR) basalts (Fig. 1a). They are also very similar to model Ni contents[13] shown by the black line in Fig. 1, which is appropriate for olivine phenocrysts of primary magmas having 8%–38% MgO and for melting of both fertile and depleted upper-mantle peridotite. Model olivines are also similar to observed olivine phenocrysts in hot Archaean komatiites from Alexo[14,15] (Fig. 1b) and other occurrences of high-temperature magmas (Methods), demonstrating restricted compositions over a broad temperature range[13].

Olivines from the Palaeocene-epoch picrites in Baffin Island and West Greenland, on the other hand, have higher Ni contents (up to 3,800 p.p.m.) than those from MORBs, Archaean komatiites, and mantle peridotite (Fig. 1c). Those with high Mg-numbers are also higher in Ni than olivines that are expected to crystallize from any partial melt of normal mantle peridotite, as indicated by Ca, Mn and Fe/Mn (Fig. 2). Primary magma major-element compositions were estimated from modelling of whole-rock lava compositions[16], but sorting of olivine phenocrysts in lava flows compromises reliable Ni estimates with this method. However, the Ni contents and Mg-numbers of the olivine phenocrysts themselves are unaffected by crystal sorting. We estimated a typical Ni content of the primary magma that is required to crystallize olivines with the observed Ni and highest Mg-numbers. The high-Ni olivines require a fertile peridotite source having ~2,360 p.p.m. Ni (Methods Summary). This is about 20% higher than normal Ni peridotite with 1,960 p.p.m. Ni and is well outside the 2σ uncertainty bounds of equation (1). The range of olivine compositions at low Mg-numbers and Ni contents is a consequence of variable olivine and clinopyroxene fractionation from primary magmas (Fig. 1c and Supplementary Information section 4).

The excess-Ni problem is not confined to Baffin Island and West Greenland. Although high-precision Ni data for olivine in intraplate occurrences are limited, we find high Ni in olivines from the Ontong Java plateau, Gorgona komatiites, and the Fernandina volcano in the Galapagos islands (Fig. 3). In all cases a peridotite source provenance is indicated by Ca, Mn and Fe/Mn (refs 13, 14) in olivine (Fig. 2 and Supplementary Information). The high Ni in Gorgona olivines is clearly seen because the sample suite includes olivines with very high Mg-numbers, close to those expected for early crystals from the primary magmas. The high Ni in olivines from the other occurrences is less clear, owing to their lower Mg-numbers. For these we modelled the Ni contents of olivines that would crystallize from primary magmas and their liquid lines of descent. Using primary magma compositions from ref. 16 with the Ni contents adjusted to those expected in melts from normal fertile peridotite containing 1,960 p.p.m. Ni (Fig. 1), the Ni contents of olivines that crystallize along liquid lines of descent are always lower than those observed in olivines from the Ontong Java plateau or the Fernandina volcano, regardless of the proportions of crystallizing olivine and clinopyroxene. Rather, all successful solutions require Ni-rich fertile peridotite sources having Ni contents up to 2,360 p.p.m.

Mantle that has been depleted by prior melting events can be high in both MgO and NiO compared with more fertile sources (see equation (1)). However, the Ni contents of partial melts of highly depleted peridotite and fertile peridotites are indistinguishable at liquid MgO $<$ 25 weight per cent[13] (Supplementary Fig. 1) and the olivines of primary melts shown in Fig. 1c should be generally applicable. Therefore, Ni-rich olivine phenocrysts cannot crystallize from melts of normal depleted peridotite.

Excess Ni in olivine has also been attributed to high temperatures and pressures of melting[17]. However, this is not consistent with olivines

---

[1]Department of Earth and Planetary Sciences, Rutgers University, 610 Taylor Road, Piscataway, New Jersey 08854-8066, USA. [2]Geological and Planetary Sciences, California Institute of Technology, Pasadena, California 91125, USA. [3]Département de Géologie, Université J. Monnet (member of PRES Université de Lyon), 23 rue P. Michelon, 42023 Saint-Etienne Cedex 2, France. [4]Laboratoire Magmas et Volcans, CNRS UMR 6524, 42023 Saint Etienne, France. [5]Department of Earth and Environment, Boston University, 685 Commonwealth Avenue, Boston, Massachusetts 02215, USA. [6]Department of Geological Sciences, University of Idaho 3022, Moscow, Idaho 83844, USA.

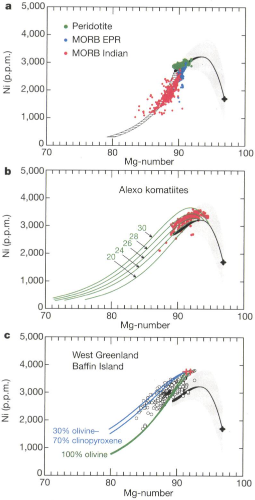


Figure 1 | Mg-numbers and Ni content for calculated[13] and observed olivine. An Mg-number is defined as 100MgO/(MgO + FeO)) in mole per cent. The black curve represents the calculated Ni content in olivines that crystallize from all primary melts derived from a fertile peridotite source having a whole-rock Ni content of 1,960 p.p.m. (ref. 13); the $\pm 1\sigma$ uncertainty (grey shading) is discussed in Supplementary Information section 3. The black hatched area represents the calculated Ni content in olivines from olivine-fractionated derivative liquids from primary magmas having 8–13% MgO. **a**, The green circles are 203 high-precision analyses of olivine in mantle peridotite from various tectonic settings (Supplementary Information); the blue and red circles are high-precision olivine phenocrysts in MORBs from the Pacific (the EPR) and Indian oceans[14,17]. **b**, The red circles are high-precision olivine phenocrysts from the Archaean Alexo komatiites[14]. The green curves are calculated Ni contents in olivines from olivine-fractionated derivative liquids from primary magmas having 20–30% MgO (ref. 13), as indicated (green numbers); primary magmas for Alexo contained 26–30% MgO (refs 13, 15). **c**, The white circles are high-precision olivine phenocrysts from West Greenland and Baffin Island[14]; the uncertainty is only slightly greater than the size of the open circles. The red crosses are the Ni contents of olivine phenocrysts expected to crystallize from primary melts derived from Ni-rich fertile peridotite having 2,360 p.p.m. Ni. The green and blue lines are olivine compositions that will crystallize from these primary magmas by fractionation of olivine and clinopyroxene in the weight proportions indicated (Supplementary Information). A low-Ni population comes from one sample (ID 27142) from West Greenland[14], and it represents 28% of the database, having Mg-number > 90. There is little overlap between this sample and the more Ni-rich types. The difference between the low- and high-Ni populations most probably reflects mantle source heterogeneity, but there are no independent geochemical data for these samples with which to test this possibility.

from Archaean komatiites from Alexo[14] (Fig. 1b) and Barberton (see Methods), which are thought to be among the hottest magmas ever erupted on Earth[15], and yet have normal Ni concentrations. Appealing to the formation of Alexo komatiites in a cool and wet environment does not solve the problem because they are compositionally similar to komatiites from the Belingwe greenstone belt[18] for which melt inclusion studies point to a hot and dry origin[19]. Furthermore, Barberton komatiites separated from a deep garnet-bearing residue that was hot and dry, and yet there is no excess Ni (Methods); low-temperature hydrous melting models are also not consistent with olivine phenocryst compositions (Methods). Consequently, the high Ni contents of the Baffin Island and West Greenland olivines are more plausibly related to composition rather than temperature-pressure effects.

Shield volcanoes from Hawaii also contain olivine phenocrysts with increased Ni (refs 14, 17), and there is general agreement that Hawaii is melting from a source that is hotter than the ambient mantle[16,17]. High Ni contents in Hawaiian olivines have been used to infer both pyroxenite melting in the source[13,14] and high temperatures and pressures of peridotite melting[17]. Comparably high mantle potential temperatures have been inferred for both Hawaii and West Greenland occurrences (1,500–1,600 °C; ref. 16) and both erupted through thick lithosphere (Supplementary Information). If increased temperature and pressure is the main mechanism for producing olivine phenocrysts with high Ni contents[17], there should be similar Ni, Mn, Fe/Mn and Ca contents in olivines from Hawaii and from Baffin Island and West Greenland. This is not observed (Supplementary Fig. 8). Hawaiian olivines are higher in Ni by ~1,000 p.p.m. and they have substantially lower calcium (Ca) and manganese (Mn) and higher Fe/Mn contents than those from West Greenland and Baffin Island. For the latter, a peridotite source provenance is clearly indicated (Fig. 2 and Supplementary Fig. 8). By contrast, low Ca and low Mn and high Fe/Mn contents for Hawaiian olivines are an expected consequence of pyroxenite source melting, owing to residuum retention of Ca in clinopyroxene and Mn in garnet (Supplementary Fig. 5). It is pyroxenite that contributes to high Ni in Hawaiian olivines[13,14], not temperature and pressure effects[17]. However, a possible role for mixed Ni-rich and Ni-poor peridotite with pyroxenite remains to be evaluated for Hawaii.

We conclude that high Ni coupled to peridotite-generated levels of Ca, Mn and Fe/Mn (Figs 1–3 and Supplementary Figs 6 and 7) point to Ni-rich peridotite sources for Baffin Island, West Greenland, Ontong Java plateau, Isla Gorgona and Fernandina in the Galapagos islands. By 'Ni-rich', we refer to Ni contents that are significantly higher than those described by equation (1), in contrast with normal-Ni peridotite sources that melt to make modern oceanic crust at mid-ocean ridges (Fig. 1a).

Current estimates of the Ni content of fertile mantle peridotite[10,11] are based on averages which can vary by about 300 p.p.m. (ref. 13); the Ni content of olivine in such peridotites obtained from open-access web sources is also large[17]. Although our estimated Ni content for Ni-rich peridotite is roughly within this range, we emphasize that this coincidence has no meaning because whole-rock and olivine Ni data are often compiled from sources of questionable accuracy[13], and there is much less uncertainty in our more recent work[12] (Fig. 1 and Supplementary Fig. 1). Random ±300 p.p.m. variations in the Ni content of peridotite can result in olivines with Ni as low as 2,200 p.p.m. (ref. 13). Such low Ni contents have never been reported from high-precision olivine analyses in unmetasomatized mantle peridotite (Fig. 1a) and, when found in olivine phenocrysts[14], other explanations are more plausible. For example, some MORB olivines can be low in Ni (Fig. 1a), but this is plausibly explained by sequestration of Ni into a residual sulphide phase[13], which can remain stable in the residue up to 15% melting[20].

Lavas exhibiting Ni excess are also associated with elevated ^{3}He/^{4}He. Picrites from Baffin Island and West Greenland have ^{3}He/^{4}He up to 50 times the atmospheric value (Ra) (ref. 21), and they also have primitive Nd and Pb isotopic compositions[3]. For the others, the maximum ^{3}He/^{4}He (Ra) is 29 for Fernandina[2] and 18 for Gorgona[22]. No helium data are currently available for lavas from the Ontong Java plateau, but they also have primitive Nd and Pb isotopic ratios[8].

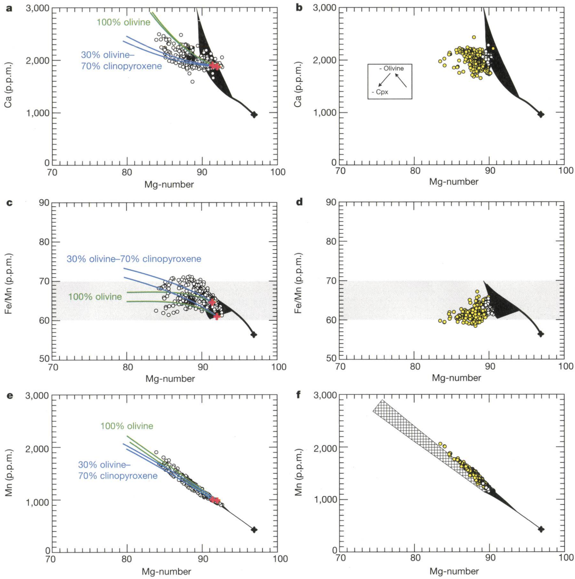


Figure 2 | Mg-numbers and Ca, Mn and Fe/Mn contents for calculated[13] and observed olivine phenocrysts from MORBs (right panels) compared with Baffin Island and West Greenland[14] (left panels). The black regions are calculated olivines from primary melts (8–38% MgO) derived by peridotite melting[13]. The black hatched area in **f** and the grey shading in **c** and **d** represent the calculated olivines from olivine-fractionated derivative liquids from primary magmas having 8–20% MgO. The green and blue lines in **a**, **c** and **e** are olivine compositions that crystallize from primary magmas that fractionated olivine and clinopyroxene in the weight proportions indicated (Supplementary Information). The inset in **b** shows how Ca in olivine changes owing to olivine and clinopyroxene fractionation from the primary magmas. The MORBs in **b**, **d** and **f** are from the EPR (white circles) and Indian Ocean (yellow circles). Although a wide range of Ca contents are theoretically possible for olivines that crystallize from peridotite-source primary magmas, most have ~2,000 p.p.m. Ca—similar to those found in MORBs, Baffin Island and West Greenland. Primitive and depleted peridotites contain ~1,000 p.p.m. Mn (refs 10, 11) and, although olivines with Fe/Mn = 55–70 can crystallize from primary magmas, more usually the range is restricted to 60–65 (ref. 13).

We now explore a model that might produce a Ni-rich peridotite with increased $^3He/^4He$. We have tested the formation of Ni-rich peridotite in a magma ocean by perovskite subtraction and/or ferropericlase addition. Although such models can yield high Ni contents, they produce a peridotite source that is too low in SiO_2 and Fe/Mn that is too high. More work on intra-mantle differentiation scenarios informed by accurate high-pressure phase equilibria and partitioning data are warranted[6], but here we propose a core–mantle interaction model.

Evidence for melting at the present-day core–mantle boundary includes the coincidence of recent determinations of the peridotite solidus with the estimated temperature at the top of the outer core[23] and seismic observations of ultralow-velocity zones[24]. As even more melt is expected in an early hotter Earth, present-day ultralow-velocity zones may represent the terminal stages in the crystallization of a long-lived magma ocean, and they provide a possible mechanism for preserving primitive mantle[5,6]. This model predicts an equilibrium exchange of Ni between liquid metal in the core and liquid silicate in the mantle, which can be described by[25–27]:

$$\mathrm{FeO}^{\mathrm{silicate}} + \mathrm{Ni}^{\mathrm{metal}} = \mathrm{NiO}^{\mathrm{silicate}} + \mathrm{Fe}^{\mathrm{metal}} \qquad (2)$$

The present-day Ni content of fertile mantle peridotite (~1,960 p.p.m.) may have been established by core formation at ~30–60 GPa (ref. 25 and references therein). Silicate melt in contact with liquid metal would be enriched in Ni at 135 GPa as the exchange is driven from left to right with increasing pressure at constant oxidation state of iron[25–27]. We propose that a Ni-rich and less-degassed domain was formed by core–mantle interaction during the late stages of the crystallization of a basal magma ocean[5] or melt layer, and that some part of this source continues to be sampled by mantle plumes. Another

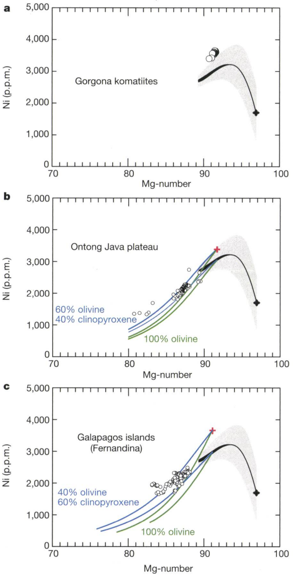


Figure 3 | Mg-numbers and Ni contents for calculated olivine[13] and observed olivine phenocrysts. **a**, Gorgona[14], **b**, the Ontong Java Plateau[14], and **c**, Fernandina (this work; Supplementary Information section 5 and Supplementary Table 2). The black curves represent the calculated Ni content in olivines of primary magmas of fertile peridotite having 1,960 p.p.m. Ni (ref. 13); the grey shaded regions ($\pm 1\sigma$) represent a possible range of Ni contents in olivine that may arise from uncertainties in the calculation (Supplementary Information). The red crosses are olivines that would crystallize from Ni-rich fertile peridotite-source primary magmas; the green and blue lines are olivine compositions that crystallize from primary magmas that fractionated olivine and clinopyroxene in the weight proportions indicated (Supplementary Information section 4).

tracer for core–mantle interaction may be the addition of tungsten and osmium (Os) isotopes from the core to the mantle[27,28]. Gorgona is the only location examined here with available ^{186}Os/^{188}Os data, and there is an ^{186}Os anomaly that has been attributed to core influence[28]. There is no shortage of plausible mechanisms of core–mantle interaction, and the challenge is to understand how they are geochemically expressed[27].

Earth's mantle may have become oxidized by disproportionation of FeO in a deep magma ocean during crystallization of Fe^{3+}-rich Mg-perovskite and loss of metallic Fe to the core[29]. The Ni-rich peridotite source for West Greenland is also oxidized, having produced melts more oxidizing than MORB and within one log unit of the nickel–nickel oxide oxygen fugacity buffer[30]. Although it is counterintuitive, we propose that these oxidized magmas had an origin that began above the core–mantle boundary, in a location where Ni-rich silicate melt and Fe^{3+}-rich Mg-perovskite were in equilibrium with the outer core. Mantle plumes that sample this region and ascend above the melting zone would solidify to Ni-rich ferropericlase, Ca perovskite and Fe^{3+}-rich Mg-perovskite, transforming further to Ni-rich and oxidized peridotite in the upper mantle. Partial melting will yield magmas like those from West Greenland that are both Ni-rich and more oxidizing than MORB.

The spatial distribution of elevated ^{3}He/^{4}He in the mantle might be constrained by understanding the origin of a Baffin Island and West Greenland source with 2,360 p.p.m. Ni. Although core formation in a single stage is not likely[26], its prediction of 3,580 p.p.m. Ni (ref. 26) in the mantle at 135 GPa may be a guide to the Ni content of silicate melt at the present time. A source with 2,360 p.p.m. Ni would then require mixing of this high-Ni endmember with the canonical ~1,960 p.p.m. Ni from elsewhere in the mantle. In this scenario, the Ni-rich and high-^{3}He/^{4}He source may be widely distributed[7,9] or it may take the form of a halo around the core. Alternatively, the isolation of the Baffin Island and West Greenland source as a refuge from convection near the core–mantle boundary[4–6,8] is plausible if it can be demonstrated that 2,360 p.p.m. Ni in a silicate melt is in equilibrium with the core. In any case, interaction of the core with a silicate melt is a plausible mechanism for production of oxidized Ni-rich peridotite, and it provides evidence for a long-lived melt layer or basal magma ocean[5,6].

METHODS SUMMARY

The method for calculating model olivine Ni compositions is given elsewhere[13] and in the Supplementary Information. Primary magma compositions are calculated first, followed by the olivines that they crystallize.

The compositions of Ni in liquids extracted from dunite [liquid + olivine] and harzburgite [liquid + olivine + orthopyroxene] residues were computed by mass-balance solutions to the equation for accumulated fractional melting:

$$C_L = C_o[1 - (1 - F)^{1/D}]/F$$

where C_L is weight per cent NiO in the liquid (primary magma), C_o is the initial NiO in the peridotite source composition (0.25%, ref. 10; 1,960 p.p.m.), F is the melt fraction, and D is the bulk distribution coefficient from the Beattie–Jones model (Supplementary Information). Ni contents for near-solidus melts were calculated from the MgO contents of near-solidus melts, together with a solidus olivine composition having 0.36% NiO (ref. 13) (Supplementary Information). Ni contents of near-solidus primary magmas are nearly indistinguishable from those of primary magmas for liquid + olivine and liquid + olivine + orthopyroxene assemblages, and the full range of possible Ni contents are shown in Supplementary Fig. 1.

The Ni content of olivine that crystallizes as a phenocryst at one atmosphere from a primary magma C_L was calculated using olivine/liquid Ni distribution coefficients from the Beattie–Jones model (Supplementary Information). Results are displayed in Figs 1 and 3 for a fertile peridotite having $C_o = 1{,}960$ p.p.m. Ni. Olivines have 2,800–3,100 p.p.m. Ni, and they provide an excellent description of olivine in normal mantle peridotite, primitive olivine phenocrysts in modern MORB, and olivine phenocrysts in Archaean komatiites. However, they fail to describe the much higher Ni contents of olivines from Baffin Island, West Greenland, the Ontong Java Plateau, Isla Gorgona and Fernandina (Galapagos) (Figs 1c and 3). For these, excellent agreement between computed and observed olivine Ni contents could be obtained with $C_o = 0.30\%$ NiO (2,360 p.p.m. Ni), a Ni-rich peridotite source.

Full Methods and any associated references are available in the online version of the paper.

Received 10 May; accepted 5 November 2012.

Published online 9 January 2013.

1. Kurz, M. D., Jenkins, W. J. & Hart, S. R. Helium isotopic systematics of oceanic islands and mantle heterogeneity. *Nature* **297**, 43–47 (1982).

2. Kurz, M. D., Curtice, J., Fornari, D., Geist, D. & Moreira, M. Primitive neon from the center of the Galàpagos hotspot. *Earth Planet. Sci. Lett.* **286,** 23–34 (2009).
3. Jackson, M. G. *et al.* Evidence for the survival of the oldest terrestrial mantle reservoir. *Nature* **466,** 853–856 (2010).
4. Tolstikhin, I. & Hofmann, A. W. Early crust on top of the Earth's core. *Phys. Earth Planet. Inter.* **148,** 109–130 (2005).
5. Labrosse, S., Hernlund, J. W. & Coltice, N. A crystallizing dense magma ocean at the base of the Earth's mantle. *Nature* **450,** 866–869 (2007).
6. Coltice, N., Moreira, M., Hernlund, J. & Labrosse, S. Crystallization of a basal magma ocean recorded by helium and neon. *Earth Planet. Sci. Lett.* **308,** 193–199 (2011).
7. Class, C. & Goldstein, S. L. Evolution of helium isotopes in the Earth's mantle. *Nature* **436,** 1107–1112 (2005).
8. Jackson, M. G. & Carlson, R. W. An ancient recipe for flood-basalt genesis. *Nature* **476,** 316–319 (2011).
9. Gonnermann, H. M. & Mukhopadhyay, S. Preserving noble gases in a convecting mantle. *Nature* **459,** 560–563 (2009).
10. McDonough, W. F. & Sun, S.-s. The composition of the Earth. *Chem. Geol.* **120,** 223–253 (1995).
11. Salters, V. J. M. & Stracke, A. Composition of depleted mantle. *Geochem. Geophys. Geosyst.* **5,** Q05B07 (2004).
12. Ionov, D. A. Compositional variations and heterogeneity in fertile lithospheric mantle: peridotite xenoliths in basalts from Tariat, Mongolia. *Contrib. Mineral. Petrol.* **154,** 455–477 (2007).
13. Herzberg, C. Identification of source lithology in the Hawaiian and Canary islands: implications for origins. *J. Petrol.* **52,** 113–146 (2011).
14. Sobolev, A. V. *et al.* The amount of recycled crust in sources of mantle-derived melts. *Science* **316,** 412–417 (2007).
15. Arndt, N. T., Lesher, C. M. & Barnes, S. J. *Komatiite* 363–389 (Cambridge Univ. Press, 2008).
16. Herzberg, C. & Gazel, E. Petrological evidence for secular cooling in mantle plumes. *Nature* **458,** 619–622 (2009).
17. Putirka, K., Ryerson, F. J., Perfit, M. & Ridley, W. I. Mineralogy and composition of the oceanic mantle. *J. Petrol.* **52,** 279–313 (2011).
18. Puchtel, I. S., Walker, R. J., Brandon, A. D. & Nisbet, E. G. Pt–Re–Os and Sm–Nd isotope and HSE and REE systematics of the 2.7 Ga Belingwe and Abitibi komatiites. *Geochim. Cosmochim. Acta* **73,** 6367–6389 (2009).
19. Berry, A. J., Danyushevsky, D. V., O'Neill, H., St C., Newville, M. & Sutton, S. R. Oxidation state of iron in komatiitic melt inclusions indicates hot Archaean mantle. *Nature* **455,** 960–963 (2008).
20. Bézos, A., Lorand, J.-P., Humler, E. & Gros, M. Platinum-group element systematic in mid-oceanic ridge basaltic glasses from the Pacific, Atlantic, and Indian oceans. *Geochim. Cosmochim. Acta* **69,** 2613–2627 (2005).
21. Starkey, N. A. *et al.* Helium isotopes in early Iceland plume picrites: constraints on the composition of high ^{3}He/^{4}He mantle. *Earth Planet. Sci. Lett.* **277,** 91–100 (2009).
22. Révillon, S. *et al.* Heterogeneity of the Caribbean plateau mantle source: new constraints from Sr, O, and He isotope compositions of olivine and clinopyroxene. *Earth Planet. Sci. Lett.* **205,** 91–106 (2002).
23. Fiquet, G. *et al.* Melting of peridotite to 140 gigapascals. *Science* **329,** 1516–1518 (2010).
24. Williams, Q. & Garnero, E. J. Seismic evidence for partial melt at the base of Earth's mantle. *Science* **273,** 1528–1530 (1996).
25. Siebert, J., Badro, J., Antonangeli, D. & Ryerson, F. J. Metal-silicate partitioning of Ni and Co in a deep magma ocean. *Earth Planet. Sci. Lett.* **321/322,** 189–197 (2012).
26. Rubie, D. *et al.* Heterogeneous accretion, composition, and core-mantle differentiation of the Earth. *Earth Planet. Sci. Lett.* **301,** 31–42 (2011).
27. Walker, D. Core-mantle chemical issues. *Can. Mineral.* **43,** 1553–1564 (2005).
28. Brandon, A. D. *et al.* ^{186}Os-^{187}Os systematic of Gorgona Island komatiites: implications for early growth of the inner core. *Earth Planet. Sci. Lett.* **206,** 411–426 (2003).
29. Wade, J. & Wood, B. J. Core formation and the oxidation state of the Earth. *Earth Planet. Sci. Lett.* **236,** 78–95 (2005).
30. Larsen, L. M. & Pedersen, A. K. Processes in high-Mg, high T magmas: evidence from olivine, chromite and glass in Paleogene picrites from West Greenland. *J. Petrol.* **41,** 1071–1098 (2000).

Supplementary Information is available in the online version of the paper.

Acknowledgements C.H. thanks L. Larsen, M. Portnyagin, A. Sobolev and D. Walker for discussions. We are very grateful to R. Walker for a critical review. D.A.I. acknowledges PNP grants from the French INSU-CNRS in 2010-2012 and P.D.A. acknowledges NSF grant EAR-1119522. P.D.A. thanks A. Matzen for extended discussions. D.G.'s work is funded by NSF grant EAR1145271.

Author Contributions C.H. modelled olivine compositions and developed a variety of magma ocean and core–mantle interaction interpretations. P.D.A. suggested the core–mantle interaction model in its current form and critiqued all Ni partition models. D.A.I. provided high-precision olivine and whole-rock analyses for mantle peridotite. C.V. acquired high-precision olivine data for Fernandina (Galapagos). M.G.J. provided information on Pb, Nd and He isotopes. D.G. provided rock samples from Fernandina. All authors contributed to the intellectual growth of this paper.

Author Information Reprints and permissions information is available at www.nature.com/reprints. The authors declare no competing financial interests. Readers are welcome to comment on the online version of the paper. Correspondence and requests for materials should be addressed to C.H. (herzberg@rci.rutgers.edu).

doi:10.1038/nature11815

Oceanographic controls on the diversity and extinction of planktonic foraminifera

Shanan E. Peters[1], Daniel C. Kelly[1] & Andrew J. Fraass[1]†

Understanding the links between long-term biological evolution, the ocean–atmosphere system and plate tectonics is a central goal of Earth science. Although environmental perturbations of many different kinds are known to have affected long-term biological evolution, particularly during major mass extinction events[1,2], the relative importance of physical environmental factors versus biological interactions in governing rates of extinction and origination through geological time remains unknown[2]. Here we use macrostratigraphic data from the Atlantic Ocean basin to show that changes in global species diversity and rates of extinction among planktonic foraminifera have been linked to tectonically and climatically forced changes in ocean circulation and chemistry from the Jurassic period to the present. Transient environmental perturbations, such as those that occurred after the asteroid impact at the end of the Cretaceous period[1] approximately 66 million years ago, and the Eocene/Oligocene greenhouse–icehouse transition[3,4] approximately 34 million years ago, are superimposed on this general long-term relationship. Rates of species origination, by contrast, are not correlated with corresponding macrostratigraphic quantities, indicating that physiochemical changes in the ocean–atmosphere system affect evolution principally by driving the synchronous extinction of lineages that originated owing to more protracted and complex interactions between biological and environmental factors.

Planktonic foraminifera are extant amoeboid protists that inhabit the surface ocean from the tropics to the poles, and their calcium carbonate tests are abundant as microfossils (<1 mm) in marine sediments globally, making them ideal tools for biostratigraphic correlation and reconstructing past environmental conditions[5]. Planktonic foraminifera also constitute a volumetrically significant source of the sediment that is deposited on oceanic crust. Approximately 55×10^6 km^2 ($\sim$67%) of the Atlantic Ocean's sea floor is blanketed by biogenic calcareous ooze[6], and the tests of planktonic foraminifera are an important constituent, contributing up to 80% of the deep-sea sedimentary calcite budget annually[7]. Previous work has attributed macroevolutionary patterns among planktonic foraminifera to several biological characteristics as well as to changes in the physical and chemical state of the ocean. For example, several lineages have independently acquired photosymbionts, an advantageous trait in oligotrophic settings[8]. The delivery of nutrients to the surface ocean and the partitioning of trophic resources between shelf and open ocean settings have also been linked to the abundance and diversification of planktonic foraminifera[9]. However, as highly effective dispersers in one of the most expansive biomes on Earth[10], the degree of partitioning of their ocean habitat through water-column stratification, latitudinal temperature gradients, sea-level fluctuations, and ocean circulation patterns is generally thought to be the most important determinant in the evolution of planktonic foraminifera[11,12].

Here we test the long-standing hypothesis that changes in the physiochemical state of the ocean have exerted a dominant influence on the long-term macroevolution of planktonic foraminifera. To do so, we compare patterns of taxonomic diversity and rates of extinction and origination to analogous macrostratigraphic quantities that describe spatiotemporal patterns of sedimentation in the deep sea. Macrostratigraphy provides a useful quantitative framework for testing this hypothesis because deep-sea sedimentation is sensitive to many important oceanographic factors, including the location of oceanic-bottom water production and its chemistry, atmospheric CO_2 concentration and the position of the calcite compensation depth (CCD), the routeing and intensity of ocean currents, and the location and extent of nutrient and solute delivery to the surface ocean[13,14]. Thus, spatiotemporal shifts in the location and extent of deep-sea sedimentation provide an integrated signal of physiochemical changes in the ocean–atmosphere system[13–16].

The total number of Atlantic drilling sites that intersect with the sea floor of a given age (Fig. 1; see Methods Summary) increases monotonically from the Jurassic period to the present (Fig. 2a). This long-term increase reflects the opening of the Atlantic Ocean basin during

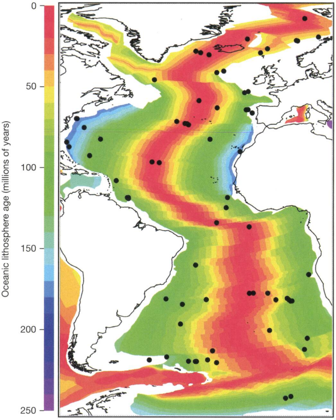


Figure 1 | Atlantic Ocean basin lithosphere ages and locations of 73 sampled drilling sites. Note that, on average, nearly 500 m of younger sediments overlie oceanic crust at each site (black dots).

---

[1]Department of Geoscience, University of Wisconsin, Madison, Wisconsin 53706, USA. †Present address: Department of Geoscience, University of Massachusetts-Amherst, Amherst, Massachusetts 01003, USA.

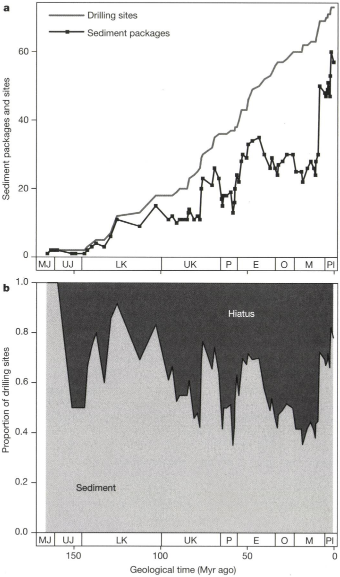


Figure 2 | Macrostratigraphy of the Atlantic Ocean basin. **a**, Total number of drilling sites intersecting the sea floor and the total number of sediment packages as a function of geologic time. **b**, Proportion of drilling sites with sediment; calculated as the total number of sites with sediment of a given age (black line in **a**) divided by the total number of sites (grey line in **a**). Myr, millions of years. Data are plotted at the age of the base of calcareous nannoplankton zones. MJ, Middle Jurassic; UJ, Upper Jurassic; LK, Lower Cretaceous; UK, Upper Cretaceous; P, Palaeocene; E, Eocene; O, Oligocene; M, Miocene; Pl, Pliocene–Pleistocene.

the Jurassic period and its ongoing expansion through sea-floor spreading. However, the total number of sites that actually preserve sediment of a given age does not increase monotonically over time owing to expansions and contractions in the proportion of sea-floor-preserving sediments. On average, ~60% of the sites in the Atlantic Ocean that could have preserved sediments at any given time actually did, but this proportion varies from less than 40% in the Palaeocene epoch to over 80% at various times during the Mesozoic era and Quaternary period (Fig. 2b).

The decline in the extent of sedimentation in the Atlantic basin from the Lower Cretaceous to the early part of the Upper Cretaceous period, and the subsequent increase to an Upper Cretaceous peak (Fig. 2), reflects shifts in the shelf-to-basin partitioning of carbonate sedimentation that have been linked to a prolonged rise and then fall in global sea level[17], a major change in the global ocean that is exemplified by the formation and draining of the North American Interior Seaway[18]. It has been postulated that the withdrawal of epicontinental seaways promoted an offshore shift in carbonate deposition during the Upper Cretaceous, with the ancillary effect of depressing the CCD[19]. Individual deep-sea records are consistent with a deepening of the CCD during this time[20], and the basin-wide increase in sedimentation during the Upper Cretaceous documented here reflects this offshore shift in carbonate burial[19].

The end of the Cretaceous period (end-Cretaceous) has long been recognized as a period of abrupt change in the ocean–atmosphere system, much of which is attributable to the Chicxulub asteroid impact[1]. The aftereffects of the end-Cretaceous crisis were prolonged[21], resulting in suppressed deep-sea sedimentation during much of the Palaeocene. This Palaeocene nadir in deep-sea sedimentation, which has been recognized globally[15], ended as sedimentation increased in extent across the Palaeocene/Eocene boundary, eventually reaching a Palaeogene maximum during the early-to-middle Eocene (Fig. 2b), possibly due to enhanced continental weathering in a globally warm climate[3]. Sedimentation in the Atlantic then declined in areal extent during the late Eocene epoch, resulting in a minimum in the early Oligocene epoch. This decline in sedimentation across the Eocene/Oligocene boundary coincides with the largest climate shift in the past 200 million years, a transition characterized by global cooling, expansion of Antarctic ice sheets, fall in sea level, and steepening of meridional temperature gradients[3,4,17].

A protracted Oligocene-to-Miocene interval of relative stasis or decline in the extent of Atlantic sedimentation was terminated by the largest expansion in sedimentation in the basin's history (Fig. 2). Like the Upper Cretaceous, the late Neogene peak in sedimentation is associated with deepening of the CCD in the Atlantic[14]. This event marks the initial development of modern-like thermohaline circulation, which involves routeing of Northern Component Water depleted in respired CO_2 towards the south and a concomitant deepening of the CCD throughout much of the Atlantic basin[22].

Macrostratigraphy quantifies spatiotemporal changes in deep-sea sedimentation that are attributable to physiochemical changes in the ocean–atmosphere system, many of which have been recognized[14–16] and identified within individual deep-sea records[23]. The independently derived global macroevolutionary history of planktonic foraminifera bears many quantitative similarities to the macrostratigraphy of the Atlantic (Fig. 3). Similarly to the number of sediment packages, there

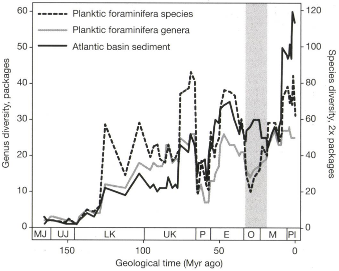


Figure 3 | Macrostratigraphy of the Atlantic Ocean basin and global range through diversity of planktonic foraminifera. De-trended time series are positively correlated (Table 1). A transient decoupling and shift to negative correlation occurred during the Oligocene–early Miocene (shaded interval). See Fig. 2 legend for abbreviations.

Table 1 | Macrostratigraphy–macroevolution correlations

	Species		Genera	
	r_s	P	r_s	P
Diversity–packages	0.526	$<1.0\times10^{-5}$	0.442	0.00011
Extinction–truncation	0.491	0.00016	0.524	0.00005
Origination–initiation	0.250	0.031	0.261	0.030

Spearman rank-order correlation coefficients (r_s) and P values between de-trended (using first differences) time series of global macroevolutionary quantities for planktonic foraminifera and corresponding macrostratigraphic quantities[30] for the Atlantic Ocean basin. To avoid edge effects and time intervals with unconstrained rates, genus-level comparisons range from NN20 to NJ12 (0.3 to 165.4 million years ago), species-level comparisons from NN20 to CC2 (0.3 to 142.4 million years ago). P values were derived from 100,000 bootstrap randomizations.

are three peaks in the global species and generic diversity of planktonic foraminifera separated by two minima, one in the Palaeocene and one in the Oligocene. The first differences in diversity, which emphasize interval-to-interval changes rather than longer-term trends, are significantly positively correlated with changes in the number of sediment packages (Table 1, correlations remain significant ($P \leq 0.006$) when the effect of interval duration is removed using partial correlation; Supplementary Table 1). However, the positive correlation between sediment packages and diversity is negative from the Oligocene to early Miocene (Fig. 3, shaded interval), weakening the overall correlations reported in Table 1. The first differences in per capita rates of extinction are also significantly positively correlated with rates of sediment-package truncation (Table 1, extinction correlations remain significant ($P \leq 0.022$) after removing the effects of interval duration; Supplementary Table 1), with the only outlier being a much larger than expected increase in species extinction in the aftermath of the end-Cretaceous asteroid impact[1]. Correlations between rates of origination and sediment-package initiation are, by contrast, weaker (Table 1) and not significant after adjusting for interval duration ($P \geq 0.09$; Supplementary Table 1).

Given that there are errors associated with both the macroevolutionary and macrostratigraphic data (see Methods), the presence of statistically significant correlations is remarkable and implies that the records are related. Comparable correlations between the sedimentary and fossil records in shallow marine and continental shelf environments have been interpreted both as a preservation-induced sampling bias in fossil data and as a shared, common signal that is induced by environmental changes associated with expansions and contractions in epicontinental seas[24].

The correlations between global macroevolutionary patterns in planktonic foraminifera and the macrostratigraphy of the Atlantic Ocean basin (Table 1) are unlikely to be due to a fossil preservation bias for several reasons. First, planktonic foraminifera are widespread geographically and only a few sites (approximately five to seven) distributed across latitudes are required to recover all extant species (see Methods). Yet, drilling sites yielding foraminifera number in the thousands globally, with almost all time intervals since the Jurassic represented by many more than seven sites. Second, the ability of Atlantic macrostratigraphy to predict global extinction better than origination (Table 1) is not consistent with the hypothesis that preservation biases have artificially clustered the first and last appearances of species at hiatuses in the sedimentary record (see Methods). Third, there is a reversal in the sign of the correlation between diversity and the number of sedimentary packages during part of the Oligocene–Miocene periods (Fig. 3), which is inconsistent with a sediment-quantity-controlled sampling bias. Fourth, the overall apparent increase in the extent of deep-sea sedimentation towards the present (Fig. 2a), a global signal attributable to sea-floor spreading and oceanic lithosphere subduction, is not mirrored by an increase in diversity, as predicted by the hypothesis that fossil preservation in sediments is driving the correlations. Last, adjusting face-value macroevolutionary rates to account for variably incomplete sampling explicitly does not qualitatively change our results. See Methods for further discussion of this optimization approach and other sampling bias related issues.

For all of these reasons, our results are best interpreted as evidence for a causal link between the processes that drive patterns of deep-sea sedimentation and the global diversity and extinction of planktonic foraminifera. Under this hypothesis, the transient decoupling of diversity and sedimentary packages that occurs at the Eocene/Oligocene boundary is attributed to a lagged macroevolutionary response to the most profound climate change of the past 200 million years[12]. Key aspects of the calcareous plankton response to the Eocene/Oligocene climatic transition are the selective extinction of warm-water taxa, equator-ward migrations of high-latitude taxa, and reduced diversities in subpolar-to-polar bioprovinces[25]. Collectively, these lines of evidence suggest that global cooling and enhanced ocean mixing reduced thermal stratification in high latitudes, thereby reducing the amount of ecospace available for segregating planktonic foraminiferal species on the basis of depth ecology[11,12]. The Oligocene increase in the extent of sedimentation, signalled by a modest upturn in sediment packages (Fig. 2), has been previously noted and attributed to a deepening of the CCD in the wake of the Eocene/Oligocene transition[26]. Thus, the reorganization of the ocean–climate system during the Eocene/Oligocene transition probably served as a common cause, driving both patterns of deep-sea sedimentation and diversity.

The ability of macrostratigraphy to predict the extinction of planktonic foraminifera better than origination is difficult to explain under the hypothesis of sampling bias, but it is readily interpretable as a biological signal. The discovery that most extant morphospecies of planktonic foraminifera consist of multiple genotypes with an extraordinary capacity to disperse, challenges the view that tectonic and hydrographic barriers facilitate speciation by isolating reproductive populations[27,28]. Instead, the genetic and zoogeographic evidence suggest that the evolution of new planktonic foraminiferal species may be limited by their ability to maintain populations in areas where unfavourable ecological conditions exist rather than by an inability to disperse beyond tectonic and hydrographic barriers[10]. An important corollary of this finding is that biotic processes such as competitive exclusion between sibling species, establishment of symbiotic relationships, differences in trophic behaviour, and divergent reproductive strategies may have a more prominent role in the origination of new species than the tectonic and climatic perturbations that consistently drive synchronous changes in the physiochemical structure of the ocean, the distribution and extent of deep-sea sedimentation, and species extinction.

METHODS SUMMARY

Data for aspects of this analysis were derived from Macrostrat (http://macrostrat.org). We compiled data on the lithology, thickness and age of sediments recovered at each of 73 offshore sites in the Atlantic Ocean (Fig. 1). Although decades of offshore drilling have resulted in globally distributed data, the Atlantic is the only basin where geographical coverage is sufficient to provide estimates of basin-scale patterns of sedimentation. Only those sites that were cored continuously from the sea floor down to or near the basaltic ocean crust were included to avoid characterizing the idiosyncratic sampling of deep-sea sediments, as opposed to their actual spatiotemporal distribution. The average modern water depth of all 73 Atlantic sites is 3.28 km, which is slightly less than the 3.92 km average depth of the Atlantic basin, indicating a modest preference for drilling on oceanic rises and in shallower water locations.

Hiatuses, which define the boundaries between otherwise continuously accumulated sediment packages, form the basis of macrostratigraphy[29,30], and were identified at each site wherever one or more calcareous nannofossil biozones (median duration 1.8 million years) were missing or not represented by a measurable thickness of sediment. A total of 249 hiatus-bound sediment packages composed of 3,692 lithologic units were recognized across all 73 Atlantic sites. Results are similar if 59 additional sites from other ocean basins, cored appropriately for macrostratigraphy, are included (see Supplementary Information).

The macroevolutionary history of planktonic foraminifera was determined by independently compiling the global stratigraphic ranges of 671 species documented in published taxonomic atlases and online resources. First and last appearance datums for planktonic foraminiferal species were binned into the same calcareous nannoplankton biozones used to recognize deep-sea hiatuses. Rates of species origination and extinction, and rates of sediment package initiation and truncation, were both calculated using per capita, per interval turnover rates[30].

Full Methods and any associated references are available in the online version of the paper.

Received 17 August; accepted 29 November 2012.

Published online 9 January 2013.

1. Schulte, P. *et al.* The Chicxulub asteroid impact and mass extinction at the Cretaceous-Paleogene boundary. *Science* **327,** 1214–1218 (2010).
2. Benton, M. J. The Red Queen and the Court Jester: species diversity and the role of biotic and abiotic factors through time. *Science* **323,** 728–732 (2009).
3. Zachos, J. C., Pagani, M., Sloan, L., Thomas, E. & Billups, K. Trends, rhythms, and aberrations in global climate 65 Ma to present. *Science* **292,** 686–693 (2001).
4. Liu, Z. *et al.* Global cooling during the Eocene-Oligocene climate transition. *Science* **323,** 1187–1190 (2009).
5. Vincent, E. & Berger, W. H. in *The Oceanic Lithosphere* (ed. Emiliani, C.) 1025–1119 (Wiley, 1981).
6. Sverdrup, H. U., Johnson, M. W. & Fleming, R. H. *The Oceans: Their Physics, Chemistry, and General Biology* (Prentice-Hall, 1942).
7. Schiebel, R. Planktonic foraminiferal sedimentation and the marine calcite budget. *Glob. Biogeochem. Cycles* **16,** 1065 (2002).
8. D'Hondt, S. & Zachos, J. C. Cretaceous foraminifera and the evolutionary history of planktonic photosymbiosis. *Paleobiology* **24,** 512–523 (1998).
9. Hallock, P., Premoli-Silva, I. & Boersma, A. Similarities between planktonic and larger foraminiferal evolutionary trends through Paleogene paleoceanographic changes. *Palaeogeogr. Palaeoclimatol. Palaeoecol.* **83,** 49–64 (1991).
10. Norris, R. D. Pelagic species diversity, biogeography, and evolution. *Paleobiology* **26,** 236–258 (2000).
11. Lipps, J. H. Plankton evolution. *Evolution* **24,** 1–22 (1970).
12. Cifelli, R. Radiation of Cenozoic planktonic foraminifera. *Syst. Zool.* **18,** 154–168 (1969).
13. Berggren, W. A. & Hollister, C. D. Plate tectonics and paleocirculation—commotion in the ocean. *Tectonophysics* **38,** 11–48 (1977).
14. Van Andel, T. H. Mesozoic/Cenozoic calcite compensation depth and the global distribution of calcareous sediments. *Earth Planet. Sci. Lett.* **26,** 187–194 (1975).
15. Moore, T. C. Jr, van Andel, T. H., Sancetta, C. & Pisias, N. Cenozoic hiatuses in pelagic sediments: marine plankton and sediments. *Micropaleontol. Sp. Pub.* **3,** 113–138 (1978).
16. Keller, G. *et al.* Global distribution of late Paleogene hiatuses. *Geology* **15,** 199–203 (1987).
17. Miller, K. G. *et al.* The Phanerozoic record of global sea-level change. *Science* **310,** 1293–1298 (2005).
18. Kauffman, E. G. & Caldwell, W. G. E. *Evolution of the Western Interior Basin* 1–30 (Geological Association of Canada Special Paper 39, 1993).
19. Milliman, J. D. in *Recent Sedimentary Carbonates* (eds Milliman, J. D., Müller, G. & Förstner, U.) Ch. 8 (Springer, 1974).
20. Thierstein, H. R. in *Deep Sea Drilling Results in the Atlantic Ocean: Continental Margins and Paleoenvironment* (eds Talwani, M., Hay, W. W. & Ryan, W. B. F.) 249–274 (American Geophysical Union, 1979).
21. D'Hondt, S., Donaghay, P., Zachos, J. C., Luttenberg, D. & Lindinger, M. Organic carbon fluxes and ecological recovery from the Cretaceous-Tertiary mass extinction. *Science* **282,** 276–279 (1998).
22. Lyle, M., Dadey, K. A. & Farrell, J. W. The late Miocene (11–8 Ma) eastern Pacific carbonate crash: evidence for reorganization of deep-water circulation by the closure of the Panama Gateway. *Proc. Ocean Drill. Program Sci. Results* **138,** 821–838 (1995).
23. Norris, R. D., Klaus, A. & Kroon, D. in *Western North Atlantic Paleogene and Cretaceous Paleoceanography* (eds Kroon, D., Norris, R. D. & Klaus, A.) 23–48 (Geological Society of London, 2001).
24. Hannisdal, B. & Peters, S. E. Phanerozoic Earth system evolution and marine biodiversity. *Science* **334,** 1121–1124 (2011).
25. Haq, B. U., Premoli-Silva, I. & Lohmann, G. P. Calcareous plankton paleobiogeographic evidence for major climatic fluctuations in the Early Cenozoic Atlantic Ocean. *J. Geophys. Res.* **82,** 3861–3876 (1977).
26. Thunell, R. C. & Corliss, B. H. in *Terminal Eocene Events* (eds Pomerol, C. & Premoli-Silva, I.) 363–380 (Elsevier, 1986).
27. Darling, K. F., Wade, C. M., Kroon, D., Brown, L. & Bijma, J. The diversity and distribution of modern planktonic foraminiferal SSU rRNA genotypes and their potential as tracers of present and past ocean circulation. *Paleoceanography* **14,** 3–12 (1999).
28. Darling, K. F. *et al.* Molecular evidence for genetic mixing of Arctic and Antarctic subpolar populations of planktonic foraminifers. *Nature* **405,** 43–47 (2000).
29. Peters, S. E. Macrostratigraphy of North America. *J. Geol.* **114,** 391–412 (2006).
30. Hannisdal, B. & Peters, S. E. On the relationship between macrostratigraphy and geological processes: quantitative information capture and sampling robustness. *J. Geol.* **118,** 111–130 (2010).

Supplementary Information is available in the online version of the paper.

Acknowledgements We thank M. Foote and B. Wilkinson for feedback on our analyses and this manuscript, and N. A. Heim for help in drafting the maps herein. This work was partially supported by National Science Foundation Division of Earth Sciences (NSF EAR) EAR 0819931 to S.E.P. and NSF EAR 0628719 subcontract to D.C.K.

Author Contributions All authors compiled Macrostrat data and contributed to development of the manuscript; D.C.K. and A.J.F. compiled the planktonic foraminifera species range data.

Author Information Reprints and permissions information is available at www.nature.com/reprints. The authors declare no competing financial interests. Readers are welcome to comment on the online version of the paper. Correspondence and requests for materials should be addressed to S.E.P. (peters@geology.wisc.edu).

doi:10.1038/nature11816

Discrete genetic modules are responsible for complex burrow evolution in *Peromyscus* mice

Jesse N. Weber[1]†, Brant K. Peterson[1,2] & Hopi E. Hoekstra[1,2]

Relative to morphological traits, we know little about how genetics influence the evolution of complex behavioural differences in nature[1]. It is unclear how the environment influences natural variation in heritable behaviour[2], and whether complex behavioural differences evolve through few genetic changes, each affecting many aspects of behaviour, or through the accumulation of several genetic changes that, when combined, give rise to behavioural complexity[3]. Here we show that in nature, oldfield mice (*Peromyscus polionotus*) build complex burrows with long entrance and escape tunnels, and that burrow length is consistent across populations, although burrow depth varies with soil composition. This burrow architecture is in contrast with the small, simple burrows of its sister species, deer mice (*P. maniculatus*). When investigated under laboratory conditions, both species recapitulate their natural burrowing behaviour. Genetic crosses between the two species reveal that the derived burrows of oldfield mice are dominant and evolved through the addition of multiple genetic changes. In burrows built by first-generation backcross mice, entrance-tunnel length and the presence of an escape tunnel can be uncoupled, suggesting that these traits are modular. Quantitative trait locus analysis also indicates that tunnel length segregates as a complex trait, affected by at least three independent genetic regions, whereas the presence of an escape tunnel is associated with only a single locus. Together, these results suggest that complex behaviours—in this case, a classic 'extended phenotype'[4]—can evolve through multiple genetic changes each affecting distinct behaviour modules.

Animal architectures, such as beehives, bird nests, spider webs, termite mounds and rodent burrows, are remarkably diverse traits that can evolve through natural selection. Despite their great diversity, these extended phenotypes have similarities: they seem to be constructed through largely unlearned motor patterns; they are often consistent within a species (or population); and, when architectures differ, these differences reflect important fitness-related functions in the wild[5]. Thus, genetic changes are predicted to contribute to the evolution of different architectures, even between closely related species; however, biologists have long questioned how genetic changes can lead to the evolution of distinct behaviours[6]. Because animal architectures can be precisely measured—similarly to morphological traits, which have been successfully dissected genetically[7,8]—these are excellent traits for the genetic analysis of behaviour.

Many species of *Peromyscus* mice build burrows, which were initially described by natural historians working in the field[9–14]. However, these burrowing behaviours can also be studied in the laboratory[15,16], and our previous work showed that burrow differences among *Peromyscus* mice have a strong genetic component[17]. In particular, *P. polionotus* and *P. maniculatus* are recently diverged and interfertile sister species that build distinct burrows. *P. polionotus* is an open-field specialist and is restricted to the southeastern United States (Fig. 1a and Supplementary Fig. 1a), whereas *P. maniculatus* is a generalist, which inhabits prairie and forest habitats across much of North America (Supplementary Fig. 1c). Whereas most *Peromyscus* species build small, single-tunnel burrows (Supplementary Fig. 1d) or no burrows at all, *P. polionotus* construct unique burrows characterized by a long entrance tunnel that leads into a nest cavity, and a secondary tunnel that emanates from the nest and terminates just below the soil surface (Fig. 1b and Supplementary Fig. 1b)[9–11]. This secondary tunnel may serve several functions[11], most notably its use as an escape tunnel when predators (for example, snakes) invade the entrance tunnel (see Supplementary Video). Overall, the complex burrow architecture in *P. polionotus* is derived[17], and probably associated with adaptation to living in an exposed, open habitat[15].

To examine natural variation in burrow complexity, we measured burrow dimensions and soil composition (that is, percentage silt, sand and clay) across the range of *P. polionotus* (Fig. 1a and Supplementary

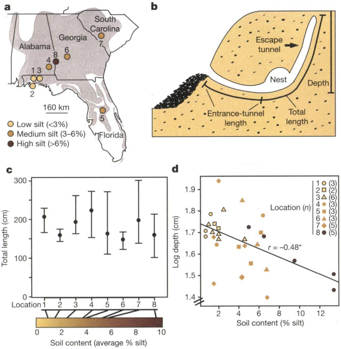


Figure 1 | Natural variation in *P. polionotus* burrows. **a**, Sampling of burrows at eight sites in the southeastern United States from across the range of *P. polionotus* (grey area). Average percentage of soil silt at each sampling site is provided. **b**, Diagram of a typical *P. polionotus* burrow showing the measures for entrance-tunnel length, total length and burrow depth, as well as a typical escape tunnel. **c**, Variation in total burrow length among sites (mean ± range; from **a**), which are ordered by increasing percentage of silt (left to right). **d**, Correlation between silt composition of soil and burrow depth (asterisk indicates Spearman correlation, rho = −0.48, $S = 6632$, $P = 0.01$). Each point represents a burrow, and shapes represent the eight different sampling sites (from **a**). The number of burrows measured at each site is shown in parentheses.

---

[1]Department of Organismic & Evolutionary Biology, Museum of Comparative Zoology, 26 Oxford Street, Cambridge, Massachusetts 02138, USA. [2]Department of Molecular & Cellular Biology, Center for Brain Science, 16 Divinity Avenue, Cambridge, Massachusetts 02138, USA. †Present address: Section of Integrative Biology, One University Station, University of Texas Austin, Texas 78712, USA.

Table 1; see Methods for details). Burrows generally had two long tunnels—an entrance and an escape tunnel (Fig. 1b)—and a mean total length of 181 cm (standard deviation (s.d.) of ± 53). Of the three soil variables measured, the percentage of silt was the most likely to influence soil compaction and thereby influence burrow dimensions[18]. Despite large differences in the percentage of silt among locations (Fig. 1c; Kruskal–Wallis (K–W) test, chi-squared = 19.42, degrees of freedom (df) = 7, $P = 0.01$), burrow length did not differ significantly among locations (K–W test, chi-squared = 6.59, df = 7, $P = 0.47$), nor did burrow length correlate with the percentage of silt measured at each burrow (Spearman test, rho = −0.16, $S = 5{,}226$, $P = 0.39$). In contrast, there was a significant, negative correlation between the percentage of silt in the soil at each burrow and burrow depth (Fig. 1d; Spearman test, rho = −0.48, $S = 6{,}632$, $P = 0.01$), suggesting that the burrows are shallower, but not shorter, when constructed in compact, silty soil[11] (Fig. 1d). Overall, natural burrow shape and length are remarkably conserved in wild *P. polionotus*, yet variation in soil composition affects burrow depth. We therefore focused our genetic analyses on burrow length.

To measure the genetic component of complex burrowing, we first assayed the burrowing behaviour of *P. polionotus* and *P. maniculatus* in a single soil type under controlled laboratory conditions (see Methods). Briefly, each assay involved placing a mouse in a large, sand-filled enclosure for 46 h (two full-night activity periods). Then, we removed the mouse and made a polyurethane cast of the burrow (Supplementary Fig. 1c, d), which we then measured (Fig. 1b). We tested each animal in three consecutive assays. Captive-reared mice, which had never been exposed to sand or allowed to burrow, recapitulated their natural burrowing behaviour in our assays. We found no significant effects of assay number, sex, age, mass or enclosure on total burrow length (data not shown). Statistical comparisons of *P. polionotus* and *P. maniculatus* burrows revealed significant differences in total burrow length (Welch's two-tailed t-test, $t = 3.24$, df = 20.98, $P < 0.01$), consistent with previous results[15,16]. Moreover, total burrow length is composed of two main parts: entrance-tunnel length, which differed significantly between species (Fig. 2b; Welch's two-tailed t-test, $t = 6.72$, df = 24.39, $P < 5 \times 10^{-3}$), and escape-tunnel length—only *P. polionotus* constructed escape tunnels (Fig. 2c). Thus, despite having no previous experience with either sand substrate or our enclosures, these species built consistently distinct burrows in the laboratory.

To determine the inheritance patterns of this behavioural variation, we characterized the genetic architecture of burrowing differences observed between *P. polionotus* and *P. maniculatus*. First, we crossed the two species and assayed the burrows of their hybrid offspring (Supplementary Methods and Fig. 2a). First-generation (F_1) hybrids built entrance tunnels that were significantly longer (after Bonferroni correction, $\alpha = 1.67 \times 10^{-2}$) than *P. maniculatus* (Welch's two-tailed t-test, $t = 5.34$, df = 20.7, $P = 2 \times 10^{-4}$). But F_1 entrance lengths did not differ from the *P. polionotus* parent (one sample t-test, $\mu = 14$ cm, $t = -0.14$, df = 12, $P = 0.89$) or from the *P. polionotus* population mean (Fig. 2b; Welch's two-tailed t-test, $t = 2.35$, df = 22.7, $P = 0.03$), and all F_1 animals constructed escape tunnels (Fig. 2c). Thus, the alleles contributing to burrow size and shape seem to segregate in a dominant fashion.

We next crossed F_1 animals with *P. maniculatus* to create a recombinant backcross (BC) generation ($n = 272$). BC mice constructed entrance tunnels that varied continuously in length between the parental extremes, but approximately one of eight ($n = 36$) of the BC mice built *P. polionotus*-length tunnels (Fig. 2b; >14 cm in length), suggesting that only a few loci are necessary to generate this behaviour. In contrast, half (46%) of the BC mice built escape tunnels (Fig. 2c). This inheritance pattern is consistent with the action of either a single major-effect locus or of multiple loci that interact to create a threshold effect[19], such that only some loci need to be co-inherited to cause the expression of a trait. Finally, tunnel lengths and the presence of escape tunnels (that is, tunnel number) are only weakly correlated in BC mice (Supplementary Table 2). Therefore, the complex burrows of *P. polionotus* comprise at least two separate behavioural modules, one for tunnel length and one for the presence of an escape tunnel.

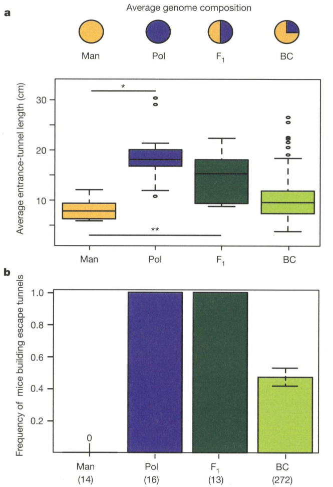


Figure 2 | Burrow variation across generations. **a**, Burrow dimensions of *P. maniculatus* (Man; yellow), *P. polionotus* (Pol; blue), F_1 hybrids (dark green) and progeny resulting from $F_1 \times$ *P. maniculatus* backcross (BC; light green). Pie charts depict average genome composition in each generation. Distributions of entrance-tunnel length (average of three trials for each individual tested) in the parental species, F_1 hybrids and BC animals are shown. Boxes represent interquartile ranges (median ± 2 s.d.). Significant t-tests, $*P = 5 \times 10^{-3}$, $**\ P = 2 \times 10^{-4}$. **b**, The frequency of escape-tunnel construction is shown for the same individuals. Error bars represent mean ± standard error of the mean (s.e.m.) Sample sizes are listed in parentheses below.

To identify the chromosomal locations involved and quantify the phenotypic effects of loci that influence these behavioural modules, we used a quantitative trait locus (QTL) mapping approach. We first identified 526 diagnostic single nucleotide polymorphisms (SNPs) using a double-digest restriction-site-associated-DNA (ddRAD) method[20], and genotyped the 272 BC mice for which we measured burrowing behaviour (that is, entrance-tunnel length and the presence of an escape tunnel; see Methods for details). Consistent with simple inheritance patterns predicted by burrow phenotypes in hybrid animals, we identified three genomic regions that contributed to variation in entrance-tunnel length, and a single region associated with escape-tunnel construction (Fig. 3a and Supplementary Fig. 5). All four QTLs are unlinked and segregate on separate chromosomes. The entrance-tunnel-length QTLs have similar effect sizes, interact additively (as determined by both a lack of epistasis between loci and graphical comparisons; Fig. 3b and Supplementary Methods), and together explain ~15% of tunnel-length variation (Table 1). Because approximately 24% of the variation in burrow length is likely to have a

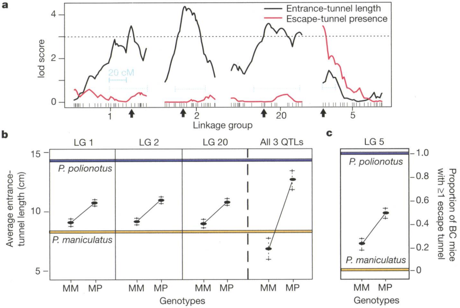


Figure 3 | QTL analysis of burrow variation. a, Linkage groups (LGs) 1, 2 and 20 harbour QTLs associated with log-transformed average entrance-tunnel length (black line). Linkage group 5 contains a single QTL associated with escape-tunnel presence (red line). Dotted line represents log odds ratio (lod) significance threshold (genome-wide $\alpha = 0.05$, lod $= \sim 3.0$). 1.5-lod confidence intervals and scale in centimorgans (cM) are shown in light blue. Dashes indicated genetic markers, and black arrows indicate markers used to define each QTL peak (used in **b**). **b,** Phenotypic effect of individual and combined QTLs (linkage groups 1, 2 and 20) on entrance-tunnel length in 272 BC mice. **c,** Proportion of BC animals that construct escape tunnels for each of the two genotypes. All error bars represent mean ± s.e.m. Blue and yellow lines represent average phenotype of the parents (pure species) used to found the cross. Genotypes are either homozygous *P. maniculatus* (MM) or heterozygous *P. maniculatus/polionotus* (MP).

genetic basis (as measured by the repeatability of burrow length in *P. polionotus*)[17], the QTLs that we report may explain more than half of the genetic variation for this trait (but see ref. 21). Even more notable, each *P. polionotus* allele increases entrance-tunnel length by, on average, 3 cm (Fig. 3b). Furthermore, BC individuals with all three *P. maniculatus* alleles versus those with one *P. polionotus* allele at each locus (that is, heterozygotes) nearly recapitulate the difference in entrance-tunnel length between the two species (that is, the mean phenotype of the parents used in the cross). Although it is possible that each QTL may harbour multiple, possibly interacting, mutations, these results are consistent with only a few genes (as few as three) being needed to explain the continuous tunnel-length variation observed in BC mice.

By contrast, QTL analysis identified only a single locus for the presence of escape tunnels on linkage group 5. Although this QTL explains only 6% of the escape-tunnel variance, BC individuals that inherited one *P. polionotus* allele at this locus were 30% more likely on average to build an escape tunnel than individuals with only *P. maniculatus* alleles (Fig. 3c). Given that the QTL is located at the end of a linkage group, it is possible that the causative mutation(s) is still far from our nearest marker. If this is the case, then we will have underestimated the phenotypic effect of this QTL, and it alone could explain the presence/absence of escape tunnels. An alternative explanation is that escape-tunnel construction is a threshold trait, but only a high-powered experiment would detect additional loci with either additive or epistatic effects on this behaviour. In either case, any additional QTLs are likely to have small phenotypic effects.

Our results show that QTLs are specific to either entrance-tunnel length or escape-tunnel presence, highlighting the modular nature of the complex *P. polionotus* burrowing behaviour. Moreover, all alleles derived from the *P. polionotus* parent cause an increase in trait value (that is, longer entrance tunnels or the addition of an escape tunnel; Fig. 3b), consistent with the role of natural selection in driving the evolution of the complex burrow architecture. Together, these results show that two aspects of burrow architecture evolved independently, as they are genetically discrete, suggesting that entrance-tunnel length and the presence of an escape tunnel are both ecologically important components of the complex *P. polionotus* burrows.

Although progress has been made towards understanding the genetic underpinnings of innate behaviours, most studies have been restricted to a few laboratory-based model systems[22] or candidate gene approaches in more diverse taxa[23]. By taking advantage of variation in an extended phenotype, in this case burrow architecture in *Peromyscus*, we have gained insight into how a seemingly complex behaviour evolves in the wild. First, we show that the complex, derived burrowing behaviour of *P. polionotus* largely results from evolution at a surprisingly small number of loci. Second, we demonstrate that burrowing behaviour is composed of distinct behavioural modules, which are controlled by independent genetic loci. This is consistent with a century-old observation that elaborate animal behaviour, such as the courtship rituals of grebes, may evolve by integrating simple behaviours[24]. Although genetic modularity has been shown to be important for the evolution of complex morphological traits[25] and laboratory-based behaviours[26], our results extend this pattern to wild

Table 1 | Models for individual and combined QTL effects

Trait	Chromosome	Position	Genotype completeness (%)	lod score	PVE	Additive effect
Log_{10} (average entrance-tunnel length)	1	75.9	77.6	2.48	3.6	0.063
	2	25.1	73.9	3.34	4.9	0.072
	20	59.6	58.8	2.47	3.6	0.064
Full entrance-tunnel-length model	-	-	-	9.55	14.8	-
Escape-tunnel presence (binary)	5	0	60.3	3.38	6.2	0.943

Full model includes all three QTLs. lod, log odds ratio; PVE, percentage of phenotypic variance explained.

behaviours. Thus, we suggest that the behavioural diversity observed in nature may often evolve by accumulating and combining alleles, each with modular effects.

METHODS SUMMARY

Experimental design. We originally obtained outbred stocks of *Peromyscus* from the *Peromyscus* Genetic Stock Center. To start the genetic cross, we mated a single female *P. maniculatus bairdii* with a male *P. polionotus subgriseus* to generate 13 F_1 hybrids, which were then backcrossed to *P. maniculatus* to produce 272 BC-generation hybrids. We assayed burrowing behaviour following our previously described methods[16]. In brief, we placed a single, virgin mouse (60–90 days of age) in an $1.2 \times 1.5 \times 1.1$ m enclosure filled with food, water, nesting material and 700 kg of hard-packed sand under constant temperature and light cycle. Each trial lasted approximately 46 h, and at the end of each trial, we made a polyurethane cast of the resulting burrow[16,27]. From each cast, we measured the length of the entrance tunnel (from the surface to the nest chamber; Fig. 1b) and recorded the presence/absence of an escape tunnel. All mice were assayed in three, consecutive 2-day trials, each time in a new, randomly assigned enclosure.

Genotyping. We extracted DNA using a high-throughput automated phenol-chloroform method (Autogen). We genotyped all individuals using a ddRAD-sequencing approach[20], which identified 526 SNPs with fixed nucleotide differences between the parental species. We used the genotypes of hybrids to estimate genetic linkage among markers. This produced a map containing 24 linkage groups, corresponding to the 24 chromosomes in *P. maniculatus* and *P. polionotus*, with a total map length of 1835.5 cM (Supplementary Fig. 2).

Data analysis. Burrow lengths were log-normalized before correlation and QTL analyses. We performed all statistical tests in the R statistical package[28]. QTLs associated with burrow phenotypes were identified using r/QTL[29]. Specifically, we sequentially performed Haley–Knot regressions, interval mapping, and interval mapping with imputation of missing genotypes, and report QTLs that are consistent across all three methods. We used permutation tests to determine genome-wide statistical significance for each QTL.

Full Methods and any associated references are available in the online version of the paper.

Received 17 August; accepted 27 November 2012.

1. Boake, C. R. B. *et al.* Genetic tools for studying adaptation and the evolution of behavior. *Am. Nat.* **160,** S143–S159 (2002).
2. West-Eberhard, M. J. *Developmental Plasticity and Evolution* (Oxford Univ. Press, 2003).
3. Mackay, T. F. C. The genetic architecture of quantitative traits. *Annu. Rev. Genet.* **35,** 303–339 (2001).
4. Dawkins, R. *The Extended Phenotype* (W. H. Freeman, 1982).
5. Hansell, M. H. *Animal Architecture* (Oxford Univ. Press, 2005).
6. Lorenz, K. Z. The evolution of behaviour. *Sci. Am.* **199,** 67–78 (1958).
7. Peichel, C. L. *et al.* The genetic architecture of divergence between threespine stickleback species. *Nature* **414,** 901–905 (2001).
8. Steiner, C. C., Weber, J. N. & Hoekstra, H. E. Adaptive variation in beach mice caused by two interacting pigmentation genes. *PLoS Biol.* **5,** 1880–1889 (2007).
9. Sumner, F. B. & Karol, J. J. Notes on the burrowing habits of *Peromyscus polionotus. J. Mamm.* **10,** 213–215 (1929).
10. Hayne, D. W. Burrowing habits of *Peromyscus polionotus. J. Mamm.* **17,** 420–421 (1936).
11. Rand, A. L. & Host, P. Mammal notes from Highland County, Florida. *Bull. Am. Mus. Nat. Hist.* **80,** 1–21 (1942).
12. Houtcooper, W. C. Rodent seed supply and burrows of *Peromyscus* in cultivated fields. *Proc. Indiana Acad. Sci.* **81,** 348–389 (1971).
13. Schwartz, C. W. & Schwartz, E. R. *The Wild Mammals of Missouri* (Univ. Missouri Press, 1981).
14. Baker, R. H. in *Biology of Peromyscus (Rodentia)* (ed. King, J. A.) (American Society of Mammalogists, 1968).
15. Wolfe, J. L. & Esher, R. J. Burrowing behaviour of old-field mice (*Peromyscus polionotus*): a laboratory investigation. *Bio. Behav.* **2,** 343–351 (1977).
16. Dawson, W. D., Lake, C. E. & Schumpert, S. S. Inheritance of burrow building in *Peromyscus. Behav. Genet.* **18,** 371–382 (1988).
17. Weber, J. N. & Hoekstra, H. E. The evolution of burrowing behavior in deer mice. *Anim. Behav.* **77,** 603–609 (2009).
18. Tan, K. H. *Soil Sampling, Preparation, and Analysis* (CRC Press, 2005).
19. Wright, S. A mutation of the guinea pig, tending to restore the pentadactyl foot when heterozygous, producing a monstrosity when homozygous. *Genetics* **20,** 84–107 (1935).
20. Peterson, B. K., Weber, J. N., Kay, E. H., Fisher, H. S. & Hoekstra, H. E. Double digest RADseq: an inexpensive method for *de novo* SNP discovery and genotyping in model and non-model species. *PLoS ONE* **7,** e37135 (2012).
21. Beavis, W. D. in *Molecular Dissection of Complex Traits* (ed. Paterson, A. H.) 431–528 (CRC Press, 1998).
22. Bendesky, A. & Bargmann, C. I. Genetic contributions to behaviour at the gene–environment interface. *Nature Rev. Genet.* **12,** 809–820 (2011).
23. Fitzpatrick, M. J. *et al.* Candidate genes for behavioural ecology. *Trends Ecol. Evol.* **20,** 96–104 (2005).
24. Huxley, J. S. The courtship habits of the great crested grebe (*Podiceps cristatus*) with an addition to the theory of sexual selection. *Proc. Zool. Soc. Lond.* **35,** 253–291 (1914).
25. Mallarino, R. *et al.* Two developmental modules establish 3D beak shape variation in Darwin's finches. *Proc. Natl Acad. Sci. USA* **108,** 4057–4062 (2011).
26. Xu, X. *et al.* Modular genetic control of sexually dimorphic behaviors. *Cell* **148,** 596–607 (2012).
27. Felthauser, M. & McInroy, D. Mapping pocket gopher burrow systems with expanding polyurethane foam. *J. Wildl. Manage.* **47,** 555–558 (1983).
28. R Development Core Team. *R: A Language and Environment for Statistical Computing* (R Foundation for Statistical Computing, 2011).
29. Broman, K. W., Wu, H., Sen, Ś. & Churchill, G. A. R/qtl: QTL mapping in experimental crosses. *Bioinformatics* **19,** 889–890 (2003).

Supplementary Information is available in the online version of the paper.

Acknowledgements We thank D. Brimmer, A. Chiu, A. Goldberg, J. Hopwood, W. Tong, S. Wolff and the Hoekstra laboratory for assistance with behavioural assays and animal husbandry; D. Haig, B. Ölveczky, N. E. Pierce and J. Sanes for helpful discussions; and Harvard's Office of Animal Resources, particularly J. Rocca and M. O'Donnell. We also thank R. Barrett, A. Bendesky, H. Fisher, E. Kay, H. Metz and W. Tong for comments on the manuscript. This research was funded by Chapman Funds for Vertebrate Locomotion to J.N.W., National Science Foundation grant (IOS-0910164) to J.N.W. and H.E.H., and an Arnold and Mabel Beckman Young Investigator Award to H.E.H.

Author Contributions J.N.W. and H.E.H. conceived and designed the experiments. B.K.P. and J.N.W. generated the ddRAD genotypes. J.N.W. performed the behaviour experiments and analysed the genetic and behavioural data. J.N.W. and H.E.H. wrote the paper.

Author Information Reprints and permissions information is available at www.nature.com/reprints. The authors declare no competing financial interests. Readers are welcome to comment on the online version of the paper. Correspondence and requests for materials should be addressed to H.E.H. (hoekstra@oeb.harvard.edu).

LETTER

doi:10.1038/nature11725

Mosaic *PPM1D* mutations are associated with predisposition to breast and ovarian cancer

Elise Ruark[1]*, Katie Snape[1]*, Peter Humburg[2]*, Chey Loveday[1], Ilirjana Bajrami[3], Rachel Brough[3,4], Daniel Nava Rodrigues[3], Anthony Renwick[1], Sheila Seal[1], Emma Ramsay[1], Silvana Del Vecchio Duarte[1], Manuel A. Rivas[2,5], Margaret Warren-Perry[1], Anna Zachariou[1], Adriana Campion-Flora[3], Sandra Hanks[1], Anne Murray[1], Naser Ansari Pour[1], Jenny Douglas[1], Lorna Gregory[2], Andrew Rimmer[2], Neil M. Walker[6], Tsun-Po Yang[7], Julian W. Adlard[8], Julian Barwell[9], Jonathan Berg[10], Angela F. Brady[11], Carole Brewer[12], Glen Brice[13], Cyril Chapman[14], Jackie Cook[15], Rosemarie Davidson[16], Alan Donaldson[17], Fiona Douglas[18], Diana Eccles[19], D. Gareth Evans[20], Lynn Greenhalgh[21], Alex Henderson[18], Louise Izatt[22], Ajith Kumar[23], Fiona Lalloo[24], Zosia Miedzybrodzka[25], Patrick J. Morrison[26], Joan Paterson[27], Mary Porteous[28], Mark T. Rogers[29], Susan Shanley[30], Lisa Walker[31], Martin Gore[32], Richard Houlston[1], Matthew A. Brown[33], Mark J. Caufield[34], Panagiotis Deloukas[7], Mark I. McCarthy[2,35,36], John A. Todd[6], The Breast and Ovarian Cancer Susceptibility Collaboration†, Wellcome Trust Case Control Consortium†, Clare Turnbull[1,30], Jorge S. Reis-Filho[3], Alan Ashworth[3], Antonis C. Antoniou[37], Christopher J. Lord[3], Peter Donnelly[2,38] & Nazneen Rahman[1,30]

Improved sequencing technologies offer unprecedented opportunities for investigating the role of rare genetic variation in common disease. However, there are considerable challenges with respect to study design, data analysis and replication[1]. Using pooled next-generation sequencing of 507 genes implicated in the repair of DNA in 1,150 samples, an analytical strategy focused on protein-truncating variants (PTVs) and a large-scale sequencing case–control replication experiment in 13,642 individuals, here we show that rare PTVs in the p53-inducible protein phosphatase *PPM1D* are associated with predisposition to breast cancer and ovarian cancer. *PPM1D* PTV mutations were present in 25 out of 7,781 cases versus 1 out of 5,861 controls ($P = 1.12 \times 10^{-5}$), including 18 mutations in 6,912 individuals with breast cancer ($P = 2.42 \times 10^{-4}$) and 12 mutations in 1,121 individuals with ovarian cancer ($P = 3.10 \times 10^{-9}$). Notably, all of the identified *PPM1D* PTVs were mosaic in lymphocyte DNA and clustered within a 370-base-pair region in the final exon of the gene, carboxy-terminal to the phosphatase catalytic domain. Functional studies demonstrate that the mutations result in enhanced suppression of p53 in response to ionizing radiation exposure, suggesting that the mutant alleles encode hyperactive PPM1D isoforms. Thus, although the mutations cause premature protein truncation, they do not result in the simple loss-of-function effect typically associated with this class of variant, but instead probably have a gain-of-function effect. Our results have implications for the detection and management of breast and ovarian cancer risk. More generally, these data provide new insights into the role of rare and of mosaic genetic variants in common conditions, and the use of sequencing in their identification.

There is strong evidence that rare genetic variation is important in breast and ovarian cancer predisposition[2,3]. In the 1990s, genome-wide linkage analysis and positional cloning led to the identification of the DNA repair genes *BRCA1* and *BRCA2*, rare mutations of which confer substantial risks of both diseases[2,3]. More recently, through case-control resequencing studies of candidate genes we, and others, have discovered rare variants that confer moderate risks of breast and/or ovarian cancer[4–10]. These cancers are therefore exemplars of the rare variant, common disease hypothesis.

The successful studies so far have focused on genes encoding proteins involved in DNA repair such as *PALB2*, *ATM*, *CHEK2*, *BRIP1*, *RAD51C* and *RAD51D* (refs 4–10). These genes are characterized by several, very rare, loss-of-function mutations, usually PTVs, which predispose carriers to breast and/or ovarian cancer[4–10]. To investigate the role of DNA repair genes in cancer susceptibility further, we sequenced 507 genes (the 'DNA repair panel') in 1,150 individuals with breast cancer from the United Kingdom, 69 of whom also had ovarian cancer (Supplementary Tables 1 and 2 and Supplementary

---

[1] Division of Genetics & Epidemiology, The Institute of Cancer Research, Sutton SM2 5NG, UK. [2] The Wellcome Trust Centre for Human Genetics, University of Oxford, Oxford OX3 7BN, UK. [3] The Breakthrough Breast Cancer Research Centre, The Institute of Cancer Research, London SW3 6JB, UK. [4] Cancer Research UK Gene Function Laboratory, The Institute of Cancer Research, London SW3 6JB, UK. [5] Nuffield Department of Clinical Medicine, University of Oxford, Oxford OX3 7LD, UK. [6] Juvenile Diabetes Research Foundation/Wellcome Trust Diabetes and Inflammation Laboratory, Cambridge Institute for Medical Research, University of Cambridge, Addenbrooke's Hospital, Cambridge CB2 0XY, UK. [7] The Wellcome Trust Sanger Institute, Wellcome Trust Genome Campus, Hinxton, Cambridge CB10 1SA, UK. [8] Yorkshire Regional Genetics Service, Chapel Allerton Hospital, Leeds LS7 4SA, UK. [9] Leicestershire Genetics Centre, University Hospitals of Leicester NHS Trust, Leicester LE1 5WW, UK. [10] Human Genetics, Division of Medical Sciences, University of Dundee, Dundee DD1 9SY, UK. [11] NW Thames Regional Genetics Service, Kennedy Galton Centre, London HA1 3UJ, UK. [12] Peninsula Regional Genetics Service, Royal Devon & Exeter Hospital, Exeter EX1 2ED, UK. [13] SW Thames Regional Genetics Service, St George's Hospital, London SW17 0RE, UK. [14] West Midlands Regional Genetics Service, Birmingham Women's Hospital, Birmingham B15 2TG, UK. [15] Sheffield Regional Genetics Service, Sheffield Children's NHS Foundation Trust, Sheffield S10 2TH, UK. [16] West of Scotland Regional Genetics Service, Laboratory Medicine, Southern General Hospital, Glasgow G51 4TF, UK. [17] South Western Regional Genetics Service, University Hospitals of Bristol NHS Foundation Trust, Bristol BS2 8EG, UK. [18] Northern Genetics Service, Newcastle upon Tyne Hospitals NHS Foundation Trust, Newcastle NE1 3BZ, UK. [19] Faculty of Medicine, University of Southampton, Southampton University Hospitals NHS Trust, Southampton SO16 5YA, UK. [20] Genetic Medicine, Manchester Academic Health Science Centre, St Mary's Hospital, Manchester M13 9WL, UK. [21] Merseyside and Cheshire Clinical Genetics Service, Liverpool Women's NHS Foundation Trust, Liverpool L8 7SS, UK. [22] SE Thames Regional Genetics Service, Guy's and St Thomas NHS Foundation Trust, London SE1 9RT, UK. [23] NE Thames Regional Genetics Service, Great Ormond St Hospital, London WC1N 3JH, UK. [24] University Department of Medical Genetics & Regional Genetics Service, St Mary's Hospital, Manchester M13 9WL, UK. [25] University of Aberdeen and North of Scotland Clinical Genetics Service, Aberdeen Royal Infirmary, Aberdeen AB25 2ZA, UK. [26] Northern Ireland Regional Genetics Service, Belfast HSC Trust, Department of Medical Genetics, Queen's University Belfast, Belfast BT9 7AB, UK. [27] East Anglian Regional Genetics Service, Cambridge University Hospitals NHS Foundation Trust, Cambridge CB2 0QQ, UK. [28] South East of Scotland Clinical Genetics Service, Western General Hospital, Edinburgh EH4 2XU, UK. [29] All Wales Medical Genetics Service, University Hospital of Wales, Cardiff CF14 4XW, UK. [30] Department of Cancer Genetics, Royal Marsden NHS Foundation Trust, Sutton SM2 5PT, UK. [31] Oxford Regional Genetics Service, Oxford University Hospitals NHS Trust, Oxford OX3 7LJ, UK. [32] Department of Gynaecologic Oncology, Royal Marsden NHS Foundation Trust, London SW3 6JJ, UK. [33] University of Queensland Diamantina Institute, University of Queensland, Princess Alexandra Hospital, Woolloongabba, Brisbane 4102, Australia. [34] Clinical Pharmacology and Barts and The London Genome Centre, William Harvey Research Institute, Barts and The London School of Medicine and Dentistry, Queen Mary University of London, London EC1M 6BQ, UK. [35] Oxford Centre for Diabetes, Endocrinology and Medicine, University of Oxford, Churchill Hospital, Oxford OX3 7LI, UK. [36] Oxford NIHR Biomedical Research Centre, Churchill Hospital, Oxford OX3 7LI, UK. [37] Centre for Cancer Genetic Epidemiology, Department of Public Health and Primary Care, University of Cambridge, Cambridge CB1 8RN, UK. [38] Department of Statistics, University of Oxford, Oxford OX1 3TG, UK.

†Lists of participants and their affiliations appear in the Supplementary Information.

*These authors contributed equally to this work.

Fig. 1). To maximize time, sample and cost efficiency we used a pooled approach combining 200 ng of DNA from each of the 24 individuals into a single pool that we hybridized to a custom pull-down containing the DNA repair panel (Supplementary Table 2). We performed sequencing using an Illumina HiSeq2000 sequencer that generated a minimum coverage per pool of 480× for at least 90% of the target region (Supplementary Fig. 2). Sequence variants were called using Syzygy[11], the performance of which was evaluated using previously generated data in a subset of the samples. The sensitivity of base substitution calling was 99.6% (439 out of 439 common variants and 24 out of 26 rare variants that were present in 1 out of 24 individuals in a pool). The sensitivity of insertion/deletion calling was 94.4% (51 out of 54 rare insertions/deletions present in 1 out of 24 individuals in a pool; Supplementary Table 3).

We next considered the 34,564 sequence variants called by Syzygy. We first focused on PTVs because of the strong association of this class of mutation with disease. In total, 1,044 PTVs were called by Syzygy and we used a PTV prioritization method to stratify the genes according to the number of different, rare truncating mutations present within the samples[12]. *PPM1D* showed the strongest signal in this analysis, and we confirmed by Sanger sequencing that five individuals carried different *PPM1D* PTVs. Two of these individuals had ovarian cancer in addition to breast cancer.

To explore the role of *PPM1D* in breast and ovarian cancer susceptibility further, we next performed a case-control Sanger sequencing analysis of *PPM1D* in a total of 13,642 individuals—7,781 unrelated individuals with breast and/or ovarian cancer and 5,861 population controls (Supplementary Table 1). We initially sequenced all *PPM1D* exons and intron–exon boundaries, but after completing this analysis in 3,803 samples we noted that all 10 PTV mutations identified occurred within the last exon of *PPM1D*, and this clustering was highly significant ($P = 8.2 \times 10^{-6}$). We thus analysed the remaining 9,839 samples for this mutation cluster region, identifying a further 16 PTVs (Supplementary Table 1 and Fig. 1). In total, we identified 25 *PPM1D* PTVs in individuals with breast and/or ovarian cancer and 1 in controls ($P = 1.12 \times 10^{-5}$; Figs 1 and 2a and Supplementary Table 4). This included 18 mutations in 6,912 individuals with breast cancer ($P = 2.42 \times 10^{-4}$) and 12 mutations in 1,121 individuals with ovarian cancer ($P = 3.10 \times 10^{-9}$). The histological features of the cancers in *PPM1D*-mutation carriers were diverse, and five individuals had both breast and ovarian cancer (Supplementary Table 5). The case series included 773 individuals with mutations in *BRCA1* or *BRCA2* (termed *BRCA1/2*-mutation carriers), four of whom also carried PTVs in *PPM1D* (4 out 773 versus 1 out of 5,861 controls, $P = 8.30 \times 10^{-4}$). We also identified a total of 16 non-synonymous, 14 synonymous and 1 intronic variant across the cases and controls; there was no evidence for an association with cancer for these variant classes (Supplementary Table 6).

The Sanger sequencing chromatograms for the *PPM1D* PTVs were unusual for heterozygous mutations as the mutant allele was considerably and consistently lower than the wild-type allele, suggesting the mutations were mosaic in lymphocyte DNA (Fig. 2a and Supplementary Fig. 3). This contrasted with the non-truncating variants

a

ID	*PPM1D* mutation	Cancer (age in yr)
1	c.1270_1363dup94	Ov ca (64), bil br ca (43, 56)
2	c.1272delGGinsC	Br ca (34)
3	c.1337C>G_p.S446X	Ov ca (43), bladder ca (55)
4	c.1340delA	Br ca (46)
5	c.1340delA	Br ca (65)
6	c.1384C>T_p.Q462X	Br ca (59)
7	c.1420delC	Ov ca (68), br ca (71)
8	c.1430delA	Br ca (44)
9	c.1434C>A_p.C478X	Br ca (40)
10	c.1448delC	Br ca (41)
11	c.1451delT	Ov ca (67)
12	c.1451delT	Bil br ca (61, 76)
13	c.1451T>G_p.L484X	Br ca (65)
14	c.1455_1456delGA	Br ca (70)
15	c.1465delT	Ov ca (60), bil br ca (50, 55)
16	c.1518delT	Ov ca (69)
17	c.1519delG	Ov ca (40), bil br ca (36, 40)
18	c.1535delA	Br ca (46)
19	c.1536insG	Ov ca (47)
20	c.1538delT	Ov ca (60), br ca (55)
21	c.1538_1551del14	Ov ca (41)
22	c.1589delC	Ov ca (69), colorectal (69)
23	c.1600_1601delTT	Br ca (62)
24	c.1613T>A_p.L538X	Br ca (63)
25	c.1637_1638dupTG	Ov ca (76)
26	c.1412delC	Control

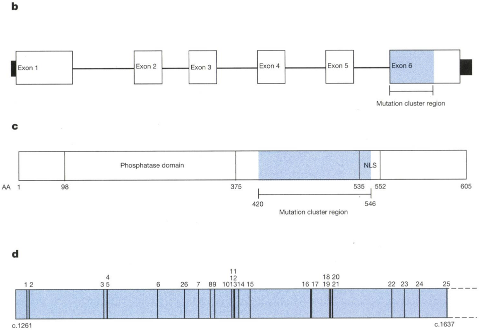


Figure 1 | Clustering of cancer predisposing mutations in *PPM1D*. **a**, *PPM1D* mutations and cancer phenotype. Bil br ca, bilateral breast cancer; br ca, breast cancer; ov ca, ovarian cancer. Age at cancer diagnosis is indicated in parentheses. **b**, *PPM1D* gene with the region targeted by mutations (mutation cluster region) in blue. **c**, PPM1D protein showing position of mutation cluster region downstream of the phosphatase domain and upstream/overlapping the nuclear localization signal (NLS). **d**, *PPM1D* mutation cluster region showing the position of mutations. The numbers above give the position of the mutations and correspond to the IDs in **a**.

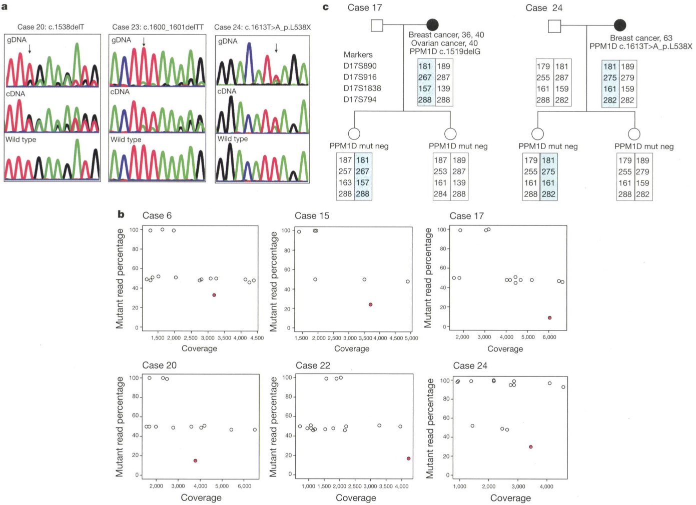


Figure 2 | *PPM1D* mutations are mosaic in lymphocyte DNA. a, Sanger sequencing traces showing mutant allele is lower in genomic DNA (gDNA) extracted from peripheral blood lymphocytes than is typical for heterozygous mutations. The cDNA analysis demonstrates that the mutations lead to truncated products, not nonsense-mediated RNA decay. **b,** Deep PCR amplicon sequencing showing heterozygous *BRCA1/2* variants at 50% (open dots), whereas the *PPM1D* mutation is present at a lower percentage (red dots). **c,** Haplotype analysis in two families. The offspring of *PPM1D*-mutation carriers have different maternal haplotypes spanning *PPM1D* (highlighted), but neither of them carries the mutation, indicating that it is either not present, or mosaic in the germline of the proband.

that all had normal sequencing profiles. DNA from saliva was available for two individuals and the PTVs were present at similar amplitude to that identified in the corresponding blood-derived DNA (Supplementary Fig. 3). To confirm that the PTV mutations were bona fide we used two other mutation detection methods: deep PCR amplicon sequencing[13] (Fig. 2b, Supplementary Fig. 4 and Supplementary Table 4) and multiplex ligation-dependent probe amplification (MLPA)[14] (Supplementary Fig. 5 and Supplementary Table 7). For the deep PCR amplicon sequencing we generated Nextera libraries of pooled PCR products covering *BRCA1*, *BRCA2* and the *PPM1D* mutation, which we sequenced using an Illumina MiSeq sequencer generating a median coverage of 3,387× across the *PPM1D* mutation (Supplementary Fig. 4 and Supplementary Table 4). This confirmed that the *PPM1D* PTVs were present at a lower proportion than heterozygous polymorphisms in *BRCA1* and *BRCA2*, with a median mutant read percentage of 16% (range 5–34%). Furthermore, we sequenced the original DNA repair panel in six cases individually (that is, unpooled), which again confirmed the mutations were present, but mosaic (Supplementary Fig. 6 and Supplementary Table 4). For three samples we had data from both the deep PCR amplicon sequencing and the DNA repair panel that gave identical mutation percentage results (Supplementary Table 4). Finally, family studies were also consistent with mosaicism; none of the 14 relatives carried the *PPM1D* mutation identified in the proband. Most compellingly, for each of probands 17 and 24, we identified two offspring that had inherited different maternal haplotypes at the *PPM1D* locus, but neither offspring carried the relevant maternal *PPM1D* mutation, demonstrating that the mutations were either not present or mosaic in the germline of the probands (Fig. 2c).

PPM1D (protein phosphatase, Mg^{2+}/Mn^{2+} dependent 1D; also known as wild-type p53-induced phosphatase 1 or *Wip1*) was first identified in a screen for p53 target genes induced by ionizing radiation[15]. *PPM1D* encodes a 605-amino-acid protein with an amino-terminal phosphatase catalytic domain and a C-terminal domain that contains a putative nuclear localization signal[16] (Fig. 1). *PPM1D* transcription is upregulated in response to various types of DNA damage in a p53-dependent manner. Once upregulated, PPM1D has been shown to dephosphorylate and downregulate several targets, particularly proteins associated with the DNA damage response (DDR) pathway initiated by the ATM (ataxia telangiectasia mutated) or ATR (ATM- and Rad3-related) protein kinases, and including tumour suppressors with a proven role in cancer susceptibility such as p53 (ref. 17), ATM[18] and CHK2 (ref. 19). Thus, it has been proposed that a primary role of PPM1D is as a homeostatic regulator of the DDR pathway, facilitating return of cells to their normal state after repair of damaged DNA[17]. There is also accumulating evidence that *PPM1D* is involved in oncogenesis[16]. *PPM1D* amplification and overexpression have been demonstrated in several human tumours[20], including breast cancers[21] and ovarian clear cell carcinoma[22], and is a promising therapeutic target[22–24].

The clustering of PTVs within the 370-base-pair (bp) region corresponding to amino acids 420–546, which is downstream of the phosphatase catalytic domain but precedes or disrupts the nuclear localization signal[25], suggests the PTVs are not acting as simple loss-of-function mutations (Fig. 1). Moreover, all the PTVs were in the last exon and thus predicted to evade nonsense-mediated RNA decay and result in a truncated protein that retains the phosphatase catalytic domain, rather than in haploinsufficiency[25,26]. We confirmed this experimentally for three mutations (Fig. 2a). To investigate the effect of *PPM1D* PTVs we generated complementary DNA expression constructs representing two mutant alleles (*PPM1D* c.1384C>T, case 6, and *PPM1D* c.1420delC, case 7) and tested their ability to suppress p53 activation in response to ionizing radiation exposure. As expected, the normal increase in p53 levels after exposure to ionizing radiation was moderately suppressed in human U2OS tumour cells transfected with a wild-type *PPM1D* expression construct, matching previous observations[16,17] (Fig. 3). The suppression of p53 was enhanced in cells transfected with the mutant *PPM1D* expression constructs, suggesting that each of these alleles encodes a hyperactive PPM1D isoform, that is, consistent with a gain-of-function rather than loss-of-function effect (Fig. 3). Similar effects were also observed in HeLa and 293 cells (Supplementary Fig. 7).

To investigate the mechanism of oncogenesis in *PPM1D* PTV-mutation carriers we analysed eight tumours from five individuals. Intriguingly, the *PPM1D* mutations were not detectable in any of the tumours by Sanger sequencing or MLPA (Supplementary Fig. 8). Through microsatellite analysis we confirmed that the tumours were from the correct individuals and demonstrated loss of heterozygosity at the *PPM1D* locus in seven out of eight tumours, although there was no evidence of *PPM1D* copy number alteration (Supplementary Fig. 8 and Supplementary Table 8). We microdissected stromal tissue from the ovarian tumour in four cases and undertook deep sequencing across the *PPM1D* PTV in blood, tumour and stromal DNA. Each mutation was present in the blood, at similar levels to that detected previously, absent from the tumour and either absent (two cases) or present at very low levels (5 out of 915 and 4 out of 5,793 reads) in the stroma, consistent with lymphocyte contamination (Supplementary Fig. 8 and Supplementary Table 5).

These data are intriguing and strongly suggest the mechanism of cancer association in *PPM1D*-mutation carriers differs from that in carriers of mutations in other DNA repair genes associated with predisposition to these cancers. There are several potential explanations. It is possible the mutation was present in the cell of cancer origin but was subsequently lost, perhaps because a *PPM1D* mutation can act as a driver to initiate oncogenesis, but is not required, or is detrimental to the progression of the resulting cancer. The allele loss we observed at the *PPM1D* locus could be interpreted as supportive of this hypothesis, but it should be noted that it is not known if the lost allele carried the *PPM1D* PTV, and loss in this region of 17q is common in these cancers. Alternatively, the absence of the *PPM1D* mutation in the tumour could be because oncogenesis is being driven by the mutation in circulating blood cells. Another possibility is that the *PPM1D* mutations are not directly involved in causing breast or ovarian cancer. For example, they could be a separate manifestation of an underlying lesion, perhaps one that causes genomic instability, which can lead to selection and clonal expansion of cells with *PPM1D* PTVs and also to cancers in other tissues. Clearly, further studies will be required to explain the mechanism of oncogenesis in *PPM1D*-mutation carriers.

Irrespective of the mechanism of the association, our data demonstrate that individuals with mosaic *PPM1D* PTVs in the mutation cluster region are at increased risk of cancer. The association is not explicable by increasing age, unlike recently reported mosaic chromosomal abnormalities[27,28] (Supplementary Table 5). To estimate the cancer risks we undertook a retrospective cohort analysis using information on breast and ovarian cancer occurrence in the 6,577 unrelated individuals negative for *BRCA1/2* mutations and controls, by modelling the retrospective likelihood of the observed mutation status conditional on the disease phenotype, as previously described[9,29]. This approach adjusts for our ascertainment of cases with more extreme phenotypes such as young age of onset or bilateral breast cancer, which we use to empower gene discovery[4–7,10,30]. The relative risk of breast cancer for *PPM1D* PTV-mutation carriers was estimated to be 2.7 (95% confidence interval: 1.3–5.3; $P = 5.38 \times 10^{-3}$), which translates to approximately 23% cumulative risk by age 80. The relative risk of ovarian cancer was estimated to be 11.5 (95% confidence interval: 4.3–30.4; $P = 9.95 \times 10^{-7}$), which translates to approximately 18% cumulative risk by age 80. It is noteworthy that we sequenced an unselected hospital-based series of 322 patients with ovarian cancer in whom we identified five *PPM1D* PTVs, suggesting that 1–2% of patients with ovarian cancer may contain mosaic *PPM1D* mutations.

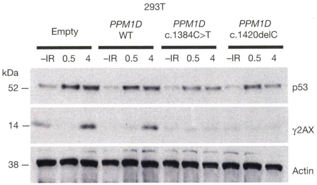


Figure 3 | The effect of mutant *PPM1D* isoforms on p53 activation. p53 wild-type U2OS human osteosarcoma cells were transfected with *PPM1D* cDNA expression constructs and exposed to ionizing radiation (IR; 5 Gy). At 30 min and 4 h intervals whole-cell lysates were analysed by western blot to estimate the ionizing-radiation-induced activation of p53. Western blots showing p53 and actin (loading control) protein levels at different times (in hours) after ionizing radiation exposure are shown. γH2AX is a marker of DNA damage, including double-stranded DNA breaks. 'Empty' denotes transfection with empty construct. PPM1D WT denotes transfection with the wild-type *PPM1D* construct. *PPM1D* c.1384C>T and *PPM1D* c.1420delC denote transfection with mutant *PPM1D* constructs. Suppression of p53 was enhanced in cells transfected with mutant constructs, suggesting that these alleles encode hyperactive PPM1D isoforms.

The frequency of *PPM1D* PTVs in *BRCA1/2*-mutation carriers with breast and/or ovarian cancer was also significantly different from population controls (4 out of 773 versus 1 out of 5,861; $P = 8.30 \times 10^{-4}$) and similar to that in cases of breast and/or ovarian cancer without *BRCA1/2* mutations (4 out of 773 versus 21 out of 6,634; $P = 0.56$), suggesting that *PPM1D* PTVs are also associated with increased risks of cancer in *BRCA1/2*-mutation carriers. Studies of unselected, population-based patients with cancer and of larger series of *BRCA1/2*-mutation carriers would be of value to extend our observations, and to explore further the prevalence and cancer risks associated with *PPM1D* mutations.

These data provide new insights into ovarian and breast cancer, potentially identifying a new class of genetic defect that lies somewhere between classic germline genetic predisposition mutations and tumour-specific somatic events. It is also highly plausible that *PPM1D* mutations are associated with other cancers, and broad evaluation of individuals with other tumour types would be of interest. More generally, the clinical implications of a mosaic cancer predisposition marker that is genetic, but not hereditary, and that is detectable in the blood but not the tumour(s) it is associated with are rather profound, particularly if this phenomenon is observed in other genes/contexts.

Our results also provide insights into genetic variation, particularly in relation to the nature and effect of rare gene mutations associated with disease. Given the truncating mutations we report probably have a gain-of-function effect, the widespread interchangeable use of the terms 'truncating mutation' and 'loss-of-function mutation' is inappropriate. We believe a more descriptive term such as 'protein-truncating variant', which does not imply the functional consequence of the mutation, is preferable. We also provide evidence that mosaic mutations can have relevance to common disease. Such variants are challenging to detect by Sanger sequencing, but are detectable by next-generation sequencing approaches. It is therefore likely that further examples of mosaic disease-associated mutations will be forthcoming, although studies to define the frequency and characteristics of mosaic mutations in control individuals will be essential, to ensure the implications of such variants in case series are correctly interpreted. Finally, although newer sequencing technologies are making large-scale whole-genome sequencing experiments ever more feasible, it is likely that focused sequencing experiments with tailored design and analytical prioritization strategies, such as those used here, will have utility over the next few years.

METHODS SUMMARY

Lymphocyte DNA from 8,046 individuals affected with breast and/or ovarian cancer and 5,861 population-based controls were included. A custom pull-down that included 507 genes (DNA repair panel) was designed using the Agilent SureSelect target enrichment system and sequenced in samples from 1,150 women, in pools of 24 samples, with an Illumina HiSeq2000. Sequence reads were mapped to the human reference genome (hg19) using BWA (version 0.5.6). Variant calling was undertaken with Syzygy[11] (version 1.2.4). Primers for *PPM1D* Sanger sequencing were designed using exon-primer. Amplicons were sequenced using the BigDye terminator cycle sequencing kit and an ABI3730 automated sequencer (ABI PerkinElmer). Deep PCR amplicon sequencing was undertaken by amplifying the *PPM1D* mutation cluster region, and *BRCA1* and *BRCA2* using a Multiplex PCR kit (Qiagen), preparing libraries with Nextera technology[11] and sequencing on an Illumina MiSeq. Sequencing of the DNA repair panel in indexed samples from six *PPM1D* PTV-mutation carriers was undertaken using Illumina TruSeq kits for library preparation and an Illumina HiSeq2000. For these latter experiments, sequence reads were mapped using Stampy version 1.0.14 and variants were called with Platypus (http://www.well.ox.ac.uk/platypus). MLPA was undertaken using the SALSA MLPA probe mix P200 (MRC Holland). Microsatellite analysis was undertaken with 5′6-FAM tagged primer pairs. PCR products were run on a 3730xL genetic analyser (ABI PerkinElmer) and data were analysed using GeneMarker v1.51 (SoftGenetics). The U2OS, HeLa and HEK293 (p53 wild-type) cell lines were transfected with a plasmid containing full-length wild-type *PPM1D* cDNA and the *PPM1D* open reading frame subcloned into pCMV6-AN-HA (Origene), generating a construct that could express a PPM1D N-terminal haemagglutinin (HA) epitope fusion protein. Mutations were introduced using the QuickChange II XL site-directed mutagenesis kit (Stratagene). Statistical analyses were performed using the statistics package in R. Cancer risks were estimated within a retrospective cohort analysis framework[9,29].

Full Methods and any associated references are available in the online version of the paper.

Received 4 May; accepted 26 October 2012.
Published online 16 December 2012.

1. Cirulli, E. T. & Goldstein, D. B. Uncovering the roles of rare variants in common disease through whole-genome sequencing. *Nature Rev. Genet.* **11,** 415–425 (2010).
2. Turnbull, C. & Rahman, N. Genetic predisposition to breast cancer: past, present, and future. *Annu. Rev. Genomics Hum. Genet.* **9,** 321–345 (2008).
3. Gayther, S. A. & Pharoah, P. D. The inherited genetics of ovarian and endometrial cancer. *Curr. Opin. Genet. Dev.* **20,** 231–238 (2010).
4. Rahman, N. *et al. PALB2*, which encodes a BRCA2-interacting protein, is a breast cancer susceptibility gene. *Nature Genet.* **39,** 165–167 (2007).
5. Renwick, A. *et al.* ATM mutations that cause ataxia-telangiectasia are breast cancer susceptibility alleles. *Nature Genet.* **38,** 873–875 (2006).
6. Seal, S. *et al.* Truncating mutations in the Fanconi anemia J gene *BRIP1* are low-penetrance breast cancer susceptibility alleles. *Nature Genet.* **38,** 1239–1241 (2006).
7. Meijers-Heijboer, H. *et al.* Low-penetrance susceptibility to breast cancer due to *CHEK2**1100delC in noncarriers of *BRCA1* or *BRCA2* mutations. *Nature Genet.* **31,** 55–59 (2002).
8. Meindl, A. *et al.* Germline mutations in breast and ovarian cancer pedigrees establish *RAD51C* as a human cancer susceptibility gene. *Nature Genet.* **42,** 410–414 (2010).
9. Loveday, C. *et al.* Germline *RAD51C* mutations confer susceptibility to ovarian cancer. *Nature Genet.* **44,** 475–476 (2012).
10. Loveday, C. *et al.* Germline mutations in *RAD51D* confer susceptibility to ovarian cancer. *Nature Genet.* **43,** 879–882 (2011).
11. Rivas, M. A. *et al.* Deep resequencing of GWAS loci identifies independent rare variants associated with inflammatory bowel disease. *Nature Genet.* **43,** 1066–1073 (2011).
12. Snape, K. *et al.* Predisposition gene identification in common cancers by exome sequencing: insights from familial breast cancer. *Breast Cancer Res. Treat.* **134,** 429–433 (2012).
13. Caruccio, N. Preparation of next-generation sequencing libraries using Nextera technology: simultaneous DNA fragmentation and adaptor tagging by *in vitro* transposition. *Methods Mol. Biol.* **733,** 241–255 (2011).
14. Schouten, J. P. *et al.* Relative quantification of 40 nucleic acid sequences by multiplex ligation-dependent probe amplification. *Nucleic Acids Res.* **30,** e57 (2002).
15. Fiscella, M. *et al.* Wip1, a novel human protein phosphatase that is induced in response to ionizing radiation in a p53-dependent manner. *Proc. Natl Acad. Sci. USA* **94,** 6048–6053 (1997).
16. Lu, X. *et al.* The type 2C phosphatase Wip1: an oncogenic regulator of tumor suppressor and DNA damage response pathways. *Cancer Metastasis Rev.* **27,** 123–135 (2008).
17. Lu, X., Nguyen, T. A. & Donehower, L. A. Reversal of the ATM/ATR-mediated DNA damage response by the oncogenic phosphatase PPM1D. *Cell Cycle* **4,** 4060–4064 (2005).
18. Shreeram, S. *et al.* Regulation of ATM/p53-dependent suppression of myc-induced lymphomas by Wip1 phosphatase. *J. Exp. Med.* **203,** 2793–2799 (2006).
19. Fujimoto, H. *et al.* Regulation of the antioncogenic Chk2 kinase by the oncogenic Wip1 phosphatase. *Cell Death Differ.* **13,** 1170–1180 (2006).
20. Bulavin, D. V. *et al.* Amplification of PPM1D in human tumors abrogates p53 tumor-suppressor activity. *Nature Genet.* **31,** 210–215 (2002).
21. Natrajan, R. *et al.* Tiling path genomic profiling of grade 3 invasive ductal breast cancers. *Clin. Cancer Res.* **15,** 2711–2722 (2009).
22. Tan, D. S. *et al.* PPM1D is a potential therapeutic target in ovarian clear cell carcinomas. *Clin. Cancer Res.* **15,** 2269–2280 (2009).
23. Rayter, S. *et al.* A chemical inhibitor of PPM1D that selectively kills cells overexpressing PPM1D. *Oncogene* **27,** 1036–1044 (2008).
24. Hayashi, R. *et al.* Optimization of a cyclic peptide inhibitor of Ser/Thr phosphatase PPM1D (Wip1). *Biochemistry* **50,** 4537–4549 (2011).
25. Chuman, Y. *et al.* PPM1D430, a novel alternative splicing variant of the human PPM1D, can dephosphorylate p53 and exhibits specific tissue expression. *J. Biochem.* **145,** 1–12 (2009).
26. Silva, A. L., Ribeiro, P., Inacio, A., Liebhaber, S. A. & Romao, L. Proximity of the poly(A)-binding protein to a premature termination codon inhibits mammalian nonsense-mediated mRNA decay. *RNA* **14,** 563–576 (2008).
27. Jacobs, K. B. *et al.* Detectable clonal mosaicism and its relationship to aging and cancer. *Nature Genet.* **44,** 651–658 (2012).
28. Laurie, C. C. *et al.* Detectable clonal mosaicism from birth to old age and its relationship to cancer. *Nature Genet.* **44,** 642–650 (2012).
29. Barnes, D. R. *et al.* Evaluation of association methods for analysing modifiers of disease risk in carriers of high risk mutations. *Genet. Epidemiol.* **36,** 274–291 (2012).
30. Antoniou, A. C. & Easton, D. F. Polygenic inheritance of breast cancer: implications for design of association studies. *Genet. Epidemiol.* **25,** 190–202 (2003).

Supplementary Information is available in the online version of the paper.

Acknowledgements We thank all the subjects and families that participated in the research and D. Dudakia, J. Bull and R. Linger for their assistance in recruitment. We are indebted to M. Stratton for discussions of the data and to A. Strydom for editorial assistance. We thank the High-Throughput Genomics Group at the Wellcome Trust Centre for Human Genetics, Oxford (funded by Wellcome Trust grant reference 090532/Z/09/Z and Medical Research Council (MRC) Hub grant G0900747 91070) for the generation of the phase 1 sequencing data. This work was funded by the Institute of Cancer Research, The Wellcome Trust, Cancer Research UK and Breakthrough Breast Cancer. We acknowledge support by the RMH-ICR National Institute for Health Research (NIHR) Specialist Biomedical Research Centre for Cancer. We acknowledge the use of DNA from the British 1958 Birth Cohort collection funded by the MRC grant G0000934 and the Wellcome Trust grant 068545/Z/02. A.C.A. is a Cancer Research UK Senior Cancer Research Fellow (C12292/A11174). P.Do. is supported by a Wolfson-Royal Society Merit Award. K.S. is supported by the Michael and Betty Kadoorie Cancer Genetics Research Programme.

Author Contributions E.R., K.S., P.H., N.M.W., T.-P.Y., M.A.B., M.J.C., C.T., J.T., M.I.M., P.De., P.Do. and N.R. (chair) are the Wellcome Trust Case Control Consortium (WTCCC) exon-resequencing group who devised and funded phase 1. J.W.A., J.Ba., J.Be., A.F.B., C.B., G.B., C.C., J.C., R.D., A.D., F.D., D.E., D.G.E., L.G., A.H., L.I., A.K., F.L., Z.M., P.J.M., J.P., M.P., M.T.R., S.Sh., L.W. and N.R. are centre leads of the Breast and Ovarian Cancer Susceptibility Collaboration (BOCS), which is coordinated by M.W.-P. and A.Z. A full list of the WTCCC and BOCS consortia is in the Supplementary Information. R.H. and M.G. assembled the unselected ovarian cancer series. L.G. coordinated phase 1 sequencing. P.H., M.R. and P.Do. undertook analysis of the pooled DNA repair panel. J.D., A.M., S.Se., S.H. and E.R. undertook NGS sequencing and analysis. S.Se. E.R., S.D.V.S., N.A.P., A.Re., K.S., C.L. and J.D. undertook Sanger sequencing, MLPA and tumour microsatellite analysis. E.R. undertook sample selection and data analyses with C.T. A.C.A. wrote the risk analysis software and oversaw the penetrance analysis that was performed by E.R. A.Ri. provided and optimized Platypus. D.N.R., A.C.-F. and J.S.R.-F. undertook histopathological analyses and performed microdissections. I.B., R.B., C.J.L. and A.A. undertook functional analyses. E.R., K.S. and N.R. managed and oversaw all aspects of the study and wrote the manuscript.

Author Information Reprints and permissions information is available at www.nature.com/reprints. The authors declare no competing financial interests. Readers are welcome to comment on the online version of the paper. Correspondence and requests for materials should be addressed to N.R. (nazneen.rahman@icr.ac.uk).

LETTER

doi:10.1038/nature11782

Exaggerated translation causes synaptic and behavioural aberrations associated with autism

Emanuela Santini[1], Thu N. Huynh[1], Andrew F. MacAskill[1], Adam G. Carter[1], Philippe Pierre[2,3,4], Davide Ruggero[5]*, Hanoch Kaphzan[1]†* & Eric Klann[1]

Autism spectrum disorders (ASDs) are an early onset, heterogeneous group of heritable neuropsychiatric disorders with symptoms that include deficits in social interaction skills, impaired communication abilities, and ritualistic-like repetitive behaviours[1,2]. One of the hypotheses for a common molecular mechanism underlying ASDs is altered translational control resulting in exaggerated protein synthesis[3]. Genetic variants in chromosome 4q, which contains the *EIF4E* locus, have been described in patients with autism[4,5]. Importantly, a rare single nucleotide polymorphism has been identified in autism that is associated with increased promoter activity in the *EIF4E* gene[6]. Here we show that genetically increasing the levels of eukaryotic translation initiation factor 4E (eIF4E) in mice[7] results in exaggerated cap-dependent translation and aberrant behaviours reminiscent of autism, including repetitive and perseverative behaviours and social interaction deficits. Moreover, these autistic-like behaviours are accompanied by synaptic pathophysiology in the medial prefrontal cortex, striatum and hippocampus. The autistic-like behaviours displayed by the eIF4E-transgenic mice are corrected by intracerebroventricular infusions of the cap-dependent translation inhibitor 4EGI-1. Our findings demonstrate a causal relationship between exaggerated cap-dependent translation, synaptic dysfunction and aberrant behaviours associated with autism.

eIF4E-transgenic mice (βT-*Eif4e*)[7] exhibited increased levels of eIF4E across brain regions (Fig. 1a) without compensatory changes in levels of other translational control proteins (Fig. 1b). We investigated whether eIF4E was bound preferentially to either eIF4E-binding protein (4E-BP) or eIF4G, which repress and promote, respectively, the initiation of cap-dependent translation[8,9]. We found significantly higher levels of eIF4E–eIF4G interactions in the brains of eIF4E-transgenic mice (Fig. 1c and Supplementary Fig. 1a), with no alterations in the interaction between eIF4E and 4E-BP (Fig. 1c, left, and Supplementary Fig. 1a). To confirm that the increased eIF4E–eIF4G interactions resulted in increased protein synthesis, we infused puromycin into the lateral ventricle of cannulated mice and labelled newly synthesized proteins using SUnSET[10,11], and observed increased *de novo* cap-dependent translation (Fig. 1d and Supplementary Fig. 1b–g). Overall, our results indicate that overexpression of eIF4E results in exaggerated cap-dependent translation in the brains of eIF4E-transgenic mice.

We then determined whether eIF4E-transgenic mice display repetitive and perseverative behaviours, which are behavioural domains required for ASD diagnosis[2]. eIF4E-transgenic mice exhibited repetitive digging behaviour in the marble-burying test[12] and increased self-grooming[13] compared with wild-type littermate controls (Fig. 2a, b). eIF4E-transgenic mice also displayed cognitive inflexibility in both a water-based Y-maze task and a modified version of the Morris water maze[14,15]. Learning ability in the acquisition and memory phases of these tasks was intact; however, in the reversal phases, eIF4E-transgenic mice were impaired in locating the new platform positions (Fig. 2c, d and Supplementary Fig. 2e–h). We tested an additional form of behavioural inflexibility by examining the eIF4E-transgenic mice for extinction of cued fear conditioning and found that they did not exhibit a significant reduction in freezing responses after extinction training (Fig. 2e). These experiments suggest that excessive cap-dependent translation in the brain affects the ability to suppress previously codified response patterns and the ability to form new behavioural strategies in response to changed environmental circumstances.

Abnormalities in social interaction skills are another behavioural defect displayed by individuals with ASDs[2]. In tests to examine social behaviour[16–18], the eIF4E-transgenic mice did not show a preference for a nonspecific stranger versus a new, inanimate object (Fig. 2f, g). Moreover, eIF4E-transgenic mice exhibited diminished reciprocal interactions with a freely moving stranger mouse (Fig. 2h), further evidence of deficits in social behaviour. The deficits in social behaviour of the eIF4E-transgenic mice are unlikely to be caused by a generalized increase in anxiety (Supplementary Fig. 2c, d, j). Moreover, the eIF4E-transgenic mice exhibited mild hyperactivity (Supplementary Fig. 2a, b), but no impairments in motor coordination, motor learning and sensorimotor gating (Supplementary Fig. 2i, k, l). Taken together, our behavioural analysis of the eIF4E-transgenic mice indicates that increased cap-dependent translation in the brain results in a distinct pattern of behavioural abnormalities consistent with ASDs.

Previous studies suggest that ASD symptoms such as cognitive inflexibility and deficits in social behaviour are generated by abnormalities in prefrontal and/or striatal circuits[19]. Consistent with this idea, the medial prefrontal cortex (PFC) is implicated in the modulation of social behaviours and social skills[20], whereas motor, social and communication impairments in boys with ASDs are associated with anatomical abnormalities in the striatum[21]. Therefore, we next examined whether the eIF4E-transgenic mice exhibited specific synaptic pathophysiologies in the medial PFC and striatum.

In the eIF4E-transgenic mice, examination of spontaneous synaptic 'mini' events in layers 2/3 of acute medial PFC slices revealed an increase in the frequency but not amplitude of excitatory events (miniature excitatory postsynaptic currents (mEPSCs); Fig. 3a), and an increase in the amplitude, but not frequency, of inhibitory events (miniature inhibitory postsynaptic currents (mIPSCs); Fig. 3b). No changes were observed in layer 5 (Supplementary Fig. 3a, b). Thus, our data suggest an enhancement of excitatory input and postsynaptic sensitivity for inhibitory events onto layer 2/3 pyramidal neurons, consistent with the hypothesis that autism may arise from an imbalance between excitatory and inhibitory synaptic transmission[22].

To determine whether the increased frequency of spontaneous mEPSCs might result from an enhanced number of synaptic contacts, we imaged dendritic spines using two-photon laser-scanning

---

[1]Center for Neural Science, New York University, New York, New York 10003, USA. [2]Centre d'Immunologie de Marseille-Luminy, Aix-Marseille Université, Marseille 13288, France. [3]INSERM, U1104, Marseille 13288, France. [4]CNRS, URM 7280, Marseille 13288, France. [5]Department of Urology, School of Medicine, Helen Diller Family Comprehensive Cancer Center, University of California, San Francisco, San Francisco, California 94143, USA. †Present address: Sagol Department of Neurobiology, Faculty of Natural Sciences, University of Haifa, Haifa 31095, Israel.

*These authors contributed equally to this work.

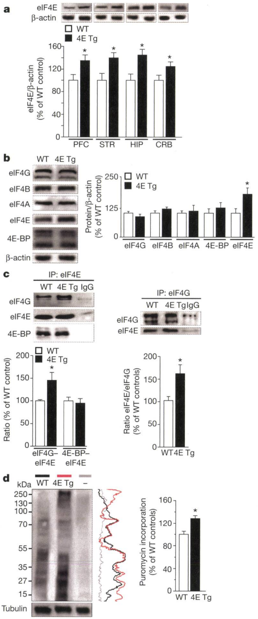


Figure 1 | eIF4E-transgenic mice exhibit increased eIF4E/eIF4G interactions and exaggerated cap-dependent translation. **a**, eIF4E-transgenic mice (4E Tg) exhibit increased eIF4E expression in multiple brain regions. $n = 4$ mice per genotype; $^*P < 0.05$ versus wild type (WT), Student's *t*-test. CRB, cerebellum; HIP, hippocampus; PFC, prefrontal cortex; STR, striatum. **b**, eIF4E-transgenic mice exhibit normal expression of other translational control proteins. $n = 4$ mice per genotype; $^*P < 0.05$, Student's *t*-test. **c**, eIF4E-transgenic mice exhibit increased eIF4E–eIF4G interactions. Immunoprecipitation (IP) of eIF4E (left) and eIF4G (right). $n = 3$ mice per genotype; $^*P < 0.05$, Student's *t*-test. **d**, eIF4E-transgenic mice exhibit increased translation as measured with SUnSET (see Methods). Vertical line traces of each autoradiogram are shown on the right. $n = 3$ mice per genotype; $^*P < 0.05$, Student's *t*-test. '–' represents a control sample without puromycin. All data are shown as mean and s.e.m.

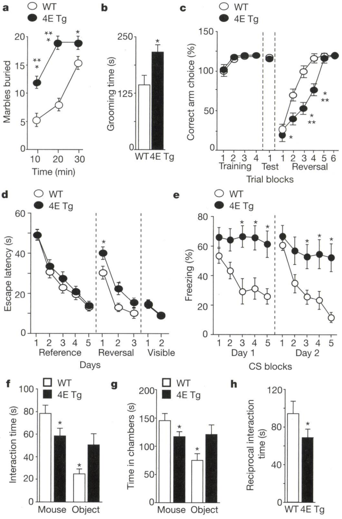


Figure 2 | eIF4E-transgenic mice exhibit ASD-like behaviours. eIF4E-transgenic mice were compared to wild-type littermates. **a**, Marble-burying test. $n = 21$–22 mice per genotype; $^*P < 0.05$, $^{***}P < 0.001$, repeated-measures analysis of variance (ANOVA) (time × genotype, $F_{(2,46)} = 31.62$, $P < 0.001$) followed by Bonferroni–Dunn post-hoc test. **b**, Self-grooming test. $n = 12$ mice per genotype; $^*P < 0.05$, Student's *t*-test. **c**, Y-maze reversal task. $n = 21$–22 mice per genotype; $^*P < 0.05$, $^{***}P < 0.001$, repeated-measures ANOVA (time × genotype, $F_{(5,138)} = 16.74$, $P < 0.001$) followed by Bonferroni–Dunn post-hoc test. **d**, Morris water maze reversal learning. $n = 12$–13 mice per genotype; $^*P < 0.05$, repeated-measures ANOVA (time × genotype, $F_{(3,92)} = 6.1$, $P < 0.001$) followed by Bonferroni–Dunn post-hoc test. **e**, Extinction of cued fear memory (15 conditioned stimuli (CS) per day represented as three CS blocks). $n = 6$ mice per genotype; $^*P < 0.05$, repeated-measures ANOVA (day 1: time × genotype, $F_{(4,40)} = 5.73$, $P < 0.001$; day 2: time × genotype, $F_{(4,40)} = 4.81$, $P < 0.01$) followed by Bonferroni–Dunn post-hoc test. **f**, **g**, Social behaviour test. The time spent either interacting with a stranger mouse (**f**) or in the chambers (**g**). $n = 6$ mice per genotype; $^*P < 0.05$, repeated-measures ANOVA (stimulus × genotype, $F_{(1,10)} = 6.04$, $P < 0.05$ (**f**); stimulus × genotype, $F_{(1,10)} = 6.12$, $P < 0.05$ (**g**)) followed by Bonferroni–Dunn post-hoc test. **h**, Reciprocal social interaction task. $n = 6$ mice per genotype; $^*P < 0.05$, Student's *t*-test. All data are shown as mean and s.e.m.

microscopy (Fig. 3c, d and Supplementary Fig. 3c, d). We found a significant increase (~12%) in spine density and observed a significantly smaller spine volume in the eIF4E-transgenic mice than in wild-type littermates (wild type = 0.123 ± 0.004 μm^3 (mean ± s.e.m.) and eIF4E-transgenic = 0.110 ± 0.004 μm^3, $P = 0.01$ versus wild type, Student's *t*-test).

Next, we examined whether increased expression of eIF4E also resulted in synaptic pathophysiology in the striatum. We used high-frequency stimulation to induce long-term depression (LTD) in acute striatal slices[23], and found that eIF4E-transgenic mice exhibited enhanced LTD compared to wild-type littermates (Fig. 3e and Supplementary Fig. 3e, f). We propose that the enhanced LTD in eIF4E-transgenic mice results in altered efficiency of striatal information storage and processing, culminating in the inability to form new motor patterns and/or to disengage from previously learned motor behaviours.

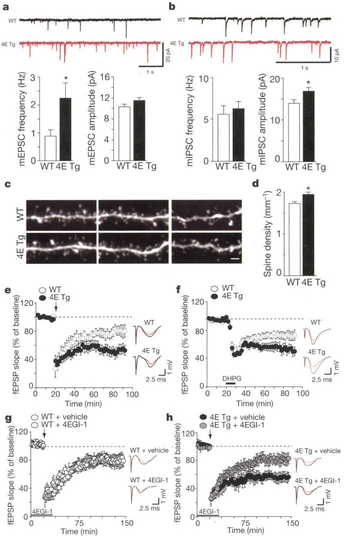


Figure 3 | eIF4E-transgenic mice exhibit alterations in synaptic function, dendritic spine density and synaptic plasticity. **a**, **b**, eIF4E-transgenic mice exhibit increased mEPSC frequency (**a**) and increased mIPSC amplitude (**b**) in layer 2/3 medial PFC pyramidal neurons. *n* = 27–30 neurons per genotype; $*P < 0.05$, Student's *t*-test. **c**, **d**, eIF4E-transgenic mice exhibit increased dendritic spine density in layer 2/3 medial PFC pyramidal neurons. High-magnification images (**c**) and quantification (**d**) of spiny dendrites. *n* = 12 neurons per genotype; $*P < 0.05$, Student's *t*-test. Scale bar, 2 µm. **e**, eIF4E-transgenic mice exhibit enhanced striatal LTD. *n* = 13 slices from 8 mice per genotype. **f**, eIF4E-transgenic mice exhibit enhanced hippocampal mGluR-LTD. DHPG denotes the mGluR agonist 3,5-dihydroxyphenylglycine. *n* = 15 slices from 8 mice per genotype. **g**, **h**, 4EGI-1 normalizes enhanced striatal LTD shown by eIF4E-transgenic mice (**h**), without affecting LTD in wild-type mice (**g**). *n* = 18 slices from 9 mice per genotype and treatment. All field recordings were analysed with repeated-measures ANOVA. Arrows indicate delivery of high-frequency stimulation. Solid bars indicate the duration of bath application of DHPG (10 µM, 10 min) and 4EGI-1 (100 µM, 45 min). Representative traces (right) showing field excitatory postsynaptic potentials (fEPSPs) before (black) and 60 min after (red) high-frequency stimulation. All data are shown as mean and s.e.m.

To determine whether the synaptic alterations described in the eIF4E-transgenic mice were selective for the frontostriatal circuit, we examined synaptic plasticity in the hippocampus[24]. We found that eIF4E-transgenic mice exhibited enhanced metabotropic glutamate receptor-dependent LTD (mGluR-LTD) compared to wild-type littermates (Fig. 3f and Supplementary Fig. 3g, h), consistent with previous studies showing that changes in brain protein synthesis are accompanied by altered (enhanced or reduced) hippocampal mGluR-LTD[25,26]. Thus, consistent with the ubiquitous increase in brain expression of eIF4E, the eIF4E-transgenic mice display altered synaptic function and plasticity in several brain regions (medial PFC, striatum and hippocampus) implicated in behavioural abnormalities associated with ASDs.

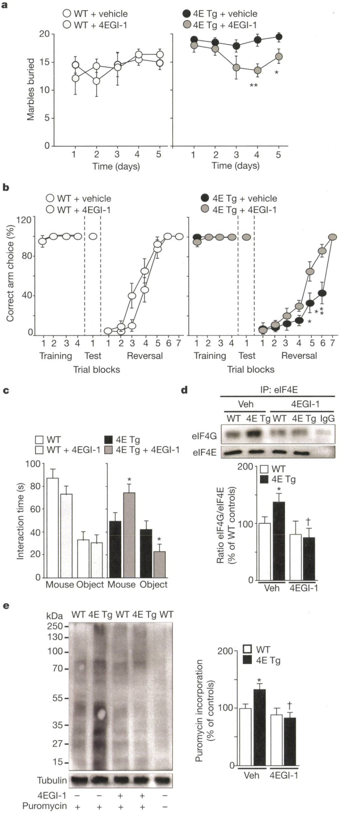


Figure 4 | The cap-dependent translation inhibitor 4EGI-1 reverses ASD-like behaviours shown by eIF4E-transgenic mice. **a**, Treatment of eIF4E-transgenic mice with 4EGI-1 reduces the marble-burying behaviour. $n = 6$ mice per genotype and treatment; $*P < 0.05$, $**P < 0.01$, two-way repeated-measures ANOVA (treatment × genotype, $F_{(1,20)} = 4.21$, $P < 0.05$) followed by Bonferroni–Dunn post-hoc test. **b**, 4EGI-1 improves the cognitive flexibility of eIF4E-transgenic mice in the Y-maze test. $n = 6$–7 mice per genotype and treatment; $*P < 0.05$, $**P < 0.01$, two-way repeated-measures ANOVA (treatment × genotype, $F_{(1,21)} = 4.61$, $P < 0.05$) followed by Bonferroni–Dunn post-hoc test. **c**, 4EGI-1 improves social behaviour of eIF4E-transgenic mice in the three-chamber arena test. $n = 6$ mice per genotype and treatment; $*P < 0.05$, two-way repeated-measures ANOVA (treatment × genotype, $F_{(1,20)} = 6.26$, $P < 0.05$) followed by Bonferroni–Dunn post-hoc test. **d**, 4EGI-1 decreases the enhanced eIF4E–eIF4G interactions in eIF4E-transgenic mice. Immunoprecipitation of eIF4E in the striatum. $n = 4$ mice per genotype; $*P < 0.05$ (versus vehicle-treated wild type), $\dagger P < 0.05$ (versus 4EGI-1-treated wild type), two-way ANOVA, followed by Bonferroni–Dunn post-hoc test. **e**, 4EGI-1 normalizes the exaggerated cap-dependent translation in eIF4E-transgenic mice as measured with SUnSET. The last wild-type sample represents a control without puromycin. $*P < 0.05$, $\dagger P < 0.05$, two-way ANOVA followed by Bonferroni–Dunn post-hoc test. All data are shown as mean and s.e.m.

Finally, we asked whether exaggerated cap-dependent translation was responsible for the synaptic alterations and ASD-like behaviours shown by the eIF4E-transgenic mice. We took advantage of 4EGI-1, an inhibitor of eIF4E–eIF4G interactions[8,11], to block the synaptic and behavioural consequences of increased eIF4E expression. Bath application of 4EGI-1 normalized the enhanced striatal LTD observed in the eIF4E-transgenic mice (Fig. 3g, h), suggesting that exaggerated striatal LTD (Fig. 3h) is a direct consequence of increased binding of eIF4E to eIF4G (Supplementary Fig. 3i–k).

Next, we used a subthreshold dose of 4EGI-1 (ref. 11) to normalize the behavioural abnormalities in eIF4E-transgenic mice without impairing their wild-type littermates. eIF4E-transgenic mice treated with 4EGI-1 exhibited a decrease in repetitive behaviour during the marble-burying task, which started on day four and persisted throughout day five (Fig. 4a). Moreover, we found that 4EGI-1 maintained the behavioural effects observed in the marble-burying task (Supplementary Fig. 4a, b). We also found that blockade of eIF4E–eIF4G interactions with 4EGI-1 significantly improved the performance of eIF4E-transgenic mice in the reversal phase of the Y-maze test (Fig. 4b). These findings indicate that chronic treatment of eIF4E-transgenic mice with 4EGI-1 reverses their repetitive and perseverative behaviours. We also found that infusions of 4EGI-1 rescued the social behaviour deficits shown by the eIF4E-transgenic mice in the three-chamber arena test, as they exhibited an increased preference for a nonspecific stranger compared to a new object (Fig. 4c).

At the completion of the behavioural studies with 4EGI-1, we performed co-immunoprecipitation experiments, confirming that 4EGI-1 reduced the increased eIF4E–eIF4G interactions exhibited by the eIF4E-transgenic mice (Fig. 4d and Supplementary Fig. 4c–e). Furthermore, puromycin-labelling of newly synthesized proteins was reduced to wild-type levels, indicating that 4EGI-1 was effective in attenuating the increased cap-dependent translation in the eIF4E-transgenic mice (Fig. 4e and Supplementary Fig. 4f, g). Together, these results indicate that repeated treatment of eIF4E-transgenic mice with 4EGI-1 reverses the increased binding of eIF4E to eIF4G, exaggerated cap-dependent translation, and reversal of ASD-like behaviours.

Here we have demonstrated that increased eIF4E expression and, consequently, dysregulated translational control at the initiation phase of protein synthesis in mice results in the appearance of synaptic dysfunction and aberrant behaviours consistent with ASDs. On the basis of our observations, we propose that exaggerated cap-dependent protein synthesis in the eIF4E-transgenic mice and fragile X syndrome model mice[27,28] results in enhanced translation of a specific subset of messenger RNAs. Thus, the identity of both these mRNAs and the *cis*-acting elements in the 5′ untranslated region responsible for eIF4E-dependent protein synthesis and their possible overlap with fragile X mental retardation protein target mRNAs will be important investigations in future studies.

Our studies with eIF4E-transgenic mice indicate that ASD-like behaviours can be induced by exaggerated cap-dependent translation in the brain. Moreover, we demonstrated that aberrant repetitive, perseverative and social behaviours shown by eIF4E-transgenic mice are reversed by reducing eIF4E–eIF4G interactions, thereby restoring translational homeostasis. Thus, our findings establish a causal link between exaggerated cap-dependent translation and behaviours associated with autism. Finally, our findings indicate that behavioural defects caused by exaggerated cap-dependent translation, which also occurs in fragile X syndrome[29,30], a disorder with a high incidence of autism, are not irrevocable and can be corrected well into adulthood.

METHODS SUMMARY

All procedures involving animals were approved by the New York University Animal Care and Use Committee and followed the National Institutes of Health Guidelines for the use of animals in research. For a detailed description of all the techniques used in this study, please see the Methods. All the experiments were performed with the examiners blinded to genotype.

Full Methods and any associated references are available in the online version of the paper.

Received 23 January; accepted 12 November 2012.

Published online 23 December 2012.

1. Levitt, P. & Campbell, D. B. The genetic and neurobiologic compass points toward common signaling dysfunctions in autism spectrum disorders. *J. Clin. Invest.* **119,** 747–754 (2009).
2. Rapin, I. & Tuchman, R. F. Autism: definition, neurobiology, screening, diagnosis. *Pediatr. Clin. North Am.* **55,** 1129–1146 (2008).
3. Kelleher, R. J. & Bear, M. F. The autistic neuron: troubled translation? *Cell* **135,** 401–406 (2008).
4. The Autism Genome Project Consortium. Mapping autism risk loci using genetic linkage and chromosomal rearrangements. *Nature Genet.* **39,** 319–328 (2007).
5. Yonan, A. L. *et al.* A genomewide screen of 345 families for autism-susceptibility loci. *Am. J. Hum. Genet.* **73,** 886–897 (2003).
6. Neves-Pereira, M. *et al.* Deregulation of *EIF4E*: a novel mechanism for autism. *J. Med. Genet.* **46,** 759–765 (2009).
7. Ruggero, D. *et al.* The translation factor eIF-4E promotes tumor formation and cooperates with c-Myc in lymphomagenesis. *Nature Med.* **10,** 484–486 (2004).
8. Moerke, N. J. *et al.* Small-molecule inhibition of the interaction between the translation initiation factors eIF4E and eIF4G. *Cell* **128,** 257–267 (2007).
9. Gingras, A. C. *et al.* Hierarchical phosphorylation of the translation inhibitor 4E-BP1. *Genes Dev.* **15,** 2852–2864 (2001).
10. Schmidt, E. K., Clavarino, G., Ceppi, M. & Pierre, P. SUnSET, a nonradioactive method to monitor protein synthesis. *Nature Methods* **6,** 275–277 (2009).
11. Hoeffer, C. A. *et al.* Inhibition of the interactions between eukaryotic initiation factors 4E and 4G impairs long-term associative memory consolidation but not reconsolidation. *Proc. Natl Acad. Sci. USA* **108,** 3383–3388 (2011).
12. Thomas, A. *et al.* Marble burying reflects a repetitive and perseverative behavior more than novelty-induced anxiety. *Psychopharmacology (Berl.)* **204,** 361–373 (2009).
13. Peça, J. *et al.* Shank3 mutant mice display autistic-like behaviours and striatal dysfunction. *Nature* **472,** 437–442 (2011).
14. Hoeffer, C. A. *et al.* Removal of FKBP12 enhances mTOR-Raptor interactions, LTP, memory, and perseverative/repetitive behavior. *Neuron* **60,** 832–845 (2008).
15. Ehninger, D. *et al.* Reversal of learning deficits in a $Tsc2^{+/-}$ mouse model of tuberous sclerosis. *Nature Med.* **14,** 843–848 (2008).
16. Moy, S. S. *et al.* Sociability and preference for social novelty in five inbred strains: an approach to assess autistic-like behavior in mice. *Genes Brain Behav.* **3,** 287–302 (2004).
17. Zhou, J. *et al.* Pharmacological inhibition of mTORC1 suppresses anatomical, cellular, and behavioral abnormalities in neural-specific Pten knock-out mice. *J. Neurosci.* **29,** 1773–1783 (2009).
18. Kwon, C.-H. *et al.* Pten regulates neuronal arborization and social interaction in mice. *Neuron* **50,** 377–388 (2006).
19. Fineberg, N. A. *et al.* Probing compulsive and impulsive behaviors, from animal models to endophenotypes: a narrative review. *Neuropsychopharmacology* **35,** 591–604 (2010).
20. Yizhar, O. *et al.* Neocortical excitation/inhibition balance in information processing and social dysfunction. *Nature* **477,** 171–178 (2011).
21. Qiu, A., Adler, M., Crocetti, D., Miller, M. I. & Mostofsky, S. H. Basal ganglia shapes predict social, communication, and motor dysfunctions in boys with autism spectrum disorder. *J. Am. Acad. Child Adolesc. Psychiatry* **49,** 539–551 (2010).
22. Tabuchi, K. *et al.* A neuroligin-3 mutation implicated in autism increases inhibitory synaptic transmission in mice. *Science* **318,** 71–76 (2007).
23. Calabresi, P., Maj, R., Pisani, A., Mercuri, N. B. & Bernardi, G. Long-term synaptic depression in the striatum: physiological and pharmacological characterization. *J. Neurosci.* **12,** 4224–4233 (1992).

24. Hou, L. *et al.* Dynamic translational and proteasomal regulation of fragile X mental retardation protein controls mGluR-dependent long-term depression. *Neuron* **51,** 441–454 (2006).
25. Huber, K. M., Kayser, M. S. & Bear, M. F. Role for rapid dendritic protein synthesis in hippocampal mGluR-dependent long-term depression. *Science* **288,** 1254–1256 (2000).
26. Auerbach, B. D., Osterweil, E. K. & Bear, M. F. Mutations causing syndromic autism define an axis of synaptic pathophysiology. *Nature* **480,** 63–68 (2011).
27. Ronesi, J. A. *et al.* Disrupted Homer scaffolds mediate abnormal mGluR5 function in a mouse model of fragile X syndrome. *Nature Neurosci.* **15,** 431–440 (2012).
28. Sharma, A. *et al.* Dysregulation of mTOR signaling in fragile X syndrome. *J. Neurosci.* **30,** 694–702 (2010).
29. Dölen, G. *et al.* Correction of fragile X syndrome in mice. *Neuron* **56,** 955–962 (2007).
30. Qin, M., Kang, J., Burlin, T. V., Jiang, C. & Smith, C. B. Postadolescent changes in regional cerebral protein synthesis: an *in vivo* study in the *Fmr1* null mouse. **25,** 5087–5095 (2005).

Supplementary Information is available in the online version of the paper.

Acknowledgements We would like to thank J. LeDoux and members of his laboratory for their technical support and suggestions. We would also like to thank D. St Clair and Z. Miedzybrodzka for their comments on the manuscript. This research was supported by National Institutes of Health (NIH) grants NS034007, NS047384 and NS078718, and Department of Defense CDMRP award W81XWH-11-1-0389 (E.K.), NIH grant CA154916 (D.R.) and the Wellcome Trust (A.F.M.).

Author Contributions The study was directed by E.K. and conceived and designed by E.S. and E.K. E.S. performed the molecular, behavioural and electrophysiological experiments. T.N.H. performed behavioural experiments. A.F.M. and A.G.C. performed the dendritic spine-density experiments. P.P. contributed the anti-puromycin (12D10) antibody. D.R. contributed with reagents and expertise concerning translation control by eIF4E. H.K. performed the cortical whole-cell electrophysiological experiments. The manuscript was written by E.S. and E.K. and edited by all of the authors.

Author Information Reprints and permissions information is available at www.nature.com/reprints. The authors declare no competing financial interests. Readers are welcome to comment on the online version of the paper. Correspondence and requests for materials should be addressed to E.K. (eklann@cns.nyu.edu).

LETTER

doi:10.1038/nature11803

Prkcz null mice show normal learning and memory

Anna M. Lee[1], Benjamin R. Kanter[1], Dan Wang[1], Jana P. Lim[1], Mimi E. Zou[1], Chichen Qiu[1], Thomas McMahon[1], Jahan Dadgar[1], Sarah C. Fischbach-Weiss[1] & Robert O. Messing[1]

Protein kinase M-ζ (PKM-ζ) is a constitutively active form of atypical protein kinase C that is exclusively expressed in the brain and implicated in the maintenance of long-term memory[1–9]. Most studies that support a role for PKM-ζ in memory maintenance have used pharmacological PKM-ζ inhibitors such as the myristoylated zeta inhibitory peptide (ZIP) or chelerythrine. Here we use a genetic approach and target exon 9 of the *Prkcz* gene to generate mice that lack both protein kinase C-ζ (PKC-ζ) and PKM-ζ (*Prkcz*$^{-/-}$ mice). *Prkcz*$^{-/-}$ mice showed normal behaviour in a cage environment and in baseline tests of motor function and sensory perception, but displayed reduced anxiety-like behaviour. Notably, *Prkcz*$^{-/-}$ mice did not show deficits in learning or memory in tests of cued fear conditioning, novel object recognition, object location recognition, conditioned place preference for cocaine, or motor learning, when compared with wild-type littermates. ZIP injection into the nucleus accumbens reduced expression of cocaine-conditioned place preference in *Prkcz*$^{-/-}$ mice. *In vitro*, ZIP and scrambled ZIP inhibited PKM-ζ, PKC-ι and PKC-ζ with similar inhibition constant (K_i) values. Chelerythrine was a weak inhibitor of PKM-ζ ($K_i = 76\,\mu$M). Our findings show that absence of PKM-ζ does not impair learning and memory in mice, and that ZIP can erase reward memory even when PKM-ζ is not present.

PKM-ζ is a constitutively active atypical kinase that is transcribed from an internal promoter in the *Prkcz* gene[10]. PKM-ζ and PKC-ζ show complementary patterns of expression, with PKM-ζ mainly expressed in the brain and PKC-ζ primarily expressed outside of the nervous system[10]. We used homologous recombination to target exon 9 of the *Prkcz* gene, which encodes the purine-binding site in the catalytic domain of PKC-ζ and PKM-ζ, to generate mice that lack both kinases. We confirmed the effect of *Prkcz* gene deletion on protein expression of PKC-ζ and PKM-ζ by western blot analysis. We detected an immunoreactive band at 52 kDa corresponding to PKM-ζ in brain samples from C57BL/6, wild-type and heterozygous *Prkcz*$^{+/-}$ mice, but not *Prkcz*$^{-/-}$ mice (Fig. 1a). A 70-kDa protein band, corresponding to PKC-ζ, was present in kidney and lung samples from wild-type mice but not from *Prkcz*$^{-/-}$ mice (Fig. 1a). These data confirm that both PKC-ζ and PKM-ζ are absent in *Prkcz*$^{-/-}$ mice.

PKC-ι is the third member of the atypical PKC subfamily that includes PKC-ζ and PKM-ζ. Although it is expressed in the brain, nothing is known about its role in regulating behaviour. Although female mice had higher levels of brain PKC-ι immunoreactivity than males, levels were similar in *Prkcz*$^{-/-}$ and wild-type mice (two-factor ANOVA, F_{sex} (1,22) = 13.56, $P = 0.001$; F_{genotype} (1,22) = 0.292, $P = 0.59$; $F_{\text{sex} \times \text{genotype}}$ (1,22) = 0.009, $P = 0.9224$; Fig. 1b, c). Because all PKC isoforms require phosphorylation at the activation loop by phosphoinositide-dependent kinase-1 (PDK1) for catalytic activity[11,12], we investigated if there was increased phosphorylation of PKC-ι at this site (T411) in *Prkcz*$^{-/-}$ mice. The ratio of phospho-T411-PKC-ι/total PKC-ι immunoreactivity (Fig. 1d) did not differ by sex or genotype (F_{sex} (1,10) = 0.096, $P = 0.76$; F_{genotype} (1,10) = 0.567, $P = 0.47$; $F_{\text{sex} \times \text{genotype}}$ (1,10) = 1.01, $P = 0.34$), and there was also no genotype difference when we combined male and female data ($t = 0.744$, $P = 0.47$). These results indicate that loss of PKM-ζ and PKC-ζ does not result in a compensatory increase in the abundance of PKC-ι or in PDK1-mediated phosphorylation of PKC-ι. We were able to detect all other PKC isozymes in brain samples except for PKC-η,

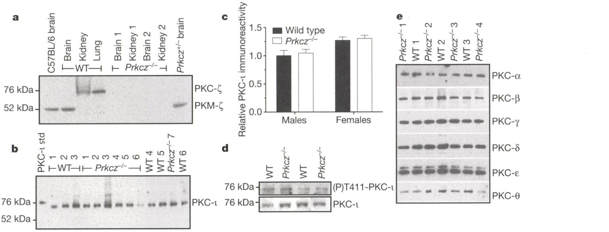


Figure 1 | Absent PKC-ζ and PKM-ζ immunoreactivity in *Prkcz*$^{-/-}$ mouse tissues. a, PKC-ζ (~72 kDa) and PKM-ζ (~52 kDa) were detected in lung, kidney and brain samples from C57BL/6, wild-type (WT) and heterozygous *Prkcz*$^{+/-}$ mice but not from *Prkcz*$^{-/-}$ mice. **b**, PKC-ι could be detected in wild-type and *Prkcz*$^{-/-}$ samples at ~72 kDa. His-tagged human PKC-ι (PKC-ι std) was run as a positive control. **c**, Females showed more brain PKC-ι immunoreactivity than males without a difference between genotypes ($n = 8$ wild-type females, $n = 6$ wild-type males, $n = 5$ *Prkcz*$^{-/-}$ females, $n = 7$ *Prkcz*$^{-/-}$ males). Data are shown as mean + s.e.m. **d**, The phospho-PKC-ι/PKC-ι ratio was similar between *Prkcz*$^{-/-}$ ($n = 7$) and wild-type mice brain samples ($n = 7$, $P = 0.47$). **e**, All PKCs, except for PKC-η, were detectable by western blot analysis in wild-type and *Prkcz*$^{-/-}$ mouse brain samples, and were of similar abundance in both genotypes.

---

[1] Ernest Gallo Clinic and Research Center, Department of Neurology, University of California, San Francisco, 5858 Horton Street, Suite 200, Emeryville, California 94608, USA.

and found that their abundance was similar in *Prkcz*$^{-/-}$ and wild-type mice (Fig. 1e and Supplementary Fig. 1).

Prkcz$^{-/-}$ mice did not show morphological abnormalities or unusual behaviours compared with wild-type mice on a standardized behavioural screen[13]. *Prkcz*$^{-/-}$ ($n = 11$) and wild-type mice displayed intact visual perception ($n = 22$) on the visual cliff test with both genotypes avoiding the perceived cliff more than 50% of the time and to a similar extent (Mann–Whitney U-test, $U = 116.5$, $P = 0.88$). There was no genotype difference in the tail-flick test for thermal nociception ($n = 13$ per genotype; $t = 0.163$, $P = 0.87$) or in total distance travelled in an open field ($n = 32$ wild type, $n = 30$ *Prkcz*$^{-/-}$ mice; $t = 0.748$, $P = 0.46$).

We analysed anxiety-like behaviour using the light–dark box and the elevated plus maze, which exploit the conflict between the desire to explore a novel environment and aversion to brightly lit, open spaces. In the light–dark box test, male *Prkcz*$^{-/-}$ mice spent 51% more time ($F_{\text{genotype} \times \text{sex}}$ (1,52) = 6.567, $P = 0.01$) and travelled 36% farther in the lit compartment ($F_{\text{genotype} \times \text{sex}}$ (1,51) = 5.803, $P = 0.02$) than male wild-type mice (Fig. 2a, b). The number of zone crossings was not different between genotype or sex (F_{genotype} (1,52) = 0.560, $P = 0.46$; F_{sex} (1,52) = 1.184, $P = 0.28$; $F_{\text{genotype} \times \text{sex}}$ (1,52) = 1.725, $P = 0.19$) (Fig. 2c). On the elevated plus maze, *Prkcz*$^{-/-}$ mice of both sexes ($n = 31$) made more entries into the open arms (F_{genotype} (1,60) = 5.615, $P = 0.02$; F_{sex} (1,60) = 0.002, $P = 0.97$; $F_{\text{genotype} \times \text{sex}}$ (1,60) = 0.013, $P = 0.91$) and tended to spend more time in the open arms (F_{genotype} (1,60) = 3.302, $P = 0.07$; F_{sex} (1,60) = 0.514, $P = 0.48$; $F_{\text{genotype} \times \text{sex}}$ (1,60) = 0.285, $P = 0.60$) than wild-type mice ($n = 33$) (Fig. 2d, e). The number of closed-arm entries was not different between genotypes or sexes (F_{genotype} (1,60) = 2.629, $P = 0.11$; F_{sex} (1,60) = 0.752, $P = 0.39$; $F_{\text{genotype} \times \text{sex}}$ (1,60) = 0.039, $P = 0.85$) (Fig. 2f). These findings indicate that *Prkcz* modulates anxiety-like behaviour, particularly in male mice.

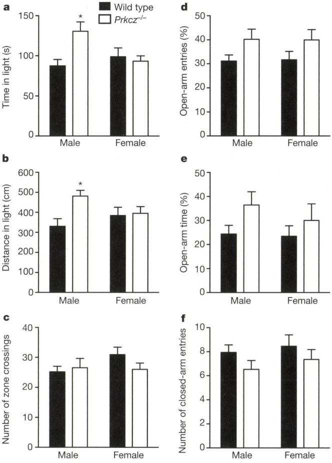


Figure 2 | Reduced anxiety-like behaviour in *Prkcz*$^{-/-}$ mice. a, b, Male *Prkcz*$^{-/-}$ mice ($n = 16$) spent more time and travelled farther in the lit compartment compared with male wild-type mice ($n = 13$). There was no genotype difference in female mice ($n = 14$ wild-type mice, $n = 13$ *Prkcz*$^{-/-}$ mice). $^*P < 0.01$ compared with male wild-type mice by Bonferroni post tests. **c**, The total number of zone crossings, a measure of locomotor activity, was similar between *Prkcz*$^{-/-}$ and wild-type mice of both sexes. **d, e,** *Prkcz*$^{-/-}$ mice ($n = 31$) made more open-arm entries and showed a trend towards spending more time in the open arms than wild-type mice ($n = 33$). **f**, The number of closed-arm entries was similar in both genotypes. Black bars represent wild-type mice; white bars represent *Prkcz*$^{-/-}$ mice; data are shown as mean + s.e.m.

Because targeting the *Prkcz* gene reduced anxiety-like behaviour, we also examined fear responses using a cued fear-conditioning procedure, by which animals learn to associate a tone with a foot shock. In this paradigm, subsequent presentations of the tone alone evoke defensive freezing behaviour. Performance in this task is impaired in rats after ZIP administration into the amygdala[7]. We first compared the response to different shock intensities and found no genotype difference (F_{genotype} (1,108) = 0.848, $P = 0.37$; F_{shock} (6,108) = 152.3, $P < 0.0001$; $F_{\text{genotype} \times \text{shock}}$ (6,108) = 1.298, $P = 0.26$). Automated scoring correlated well with hand scoring for the first tone presentation on the training day ($r^2 = 0.863$, $P < 0.0001$; Supplementary Fig. 2a), the average of five tone presentations on test day 1 ($r^2 = 0.894$; $P < 0.0001$; Supplementary Fig. 2b) and the average of three tone presentations on test day 2 ($r^2 = 0.786$; $P < 0.0003$; Supplementary Fig. 2c). There was a low level of freezing during the first tone presentation on the training day when the mice had not yet been exposed to the shock (Fig. 3a). On test day 1, freezing to the tone was significantly greater, indicating that mice had learned to associate the tone with the foot shock. The mice exhibited a similar level of freezing on test day 2. We found no genotype difference in freezing during any of the sessions, indicating that *Prkcz*$^{-/-}$ and wild-type mice learned the association equally well (F_{genotype} (1,90) = 0.070, $P = 0.79$; F_{session} (2,90) = 134.6, $P < 0.0001$; $F_{\text{genotype} \times \text{session}}$ (2,90) = 0.30, $P = 0.745$).

Given that fear memory was unimpaired in *Prkcz*$^{-/-}$ mice, we investigated other tests of learning and memory. We first used a novel object recognition task to examine hippocampal-dependent learning and memory[14]. Wild-type ($t = 3.69$, $P = 0.002$) and *Prkcz*$^{-/-}$ mice ($t = 2.98$, $P = 0.01$) showed greater exploration of the novel object compared with chance, and there was no genotype difference in time exploring the novel object ($t = 0.31$, $P = 0.76$) (Fig. 3b). We tested object location memory using a procedure in which performance is impaired in rats administered ZIP bilaterally into the hippocampus[1]. Mice of both genotypes spent more time exploring the new location (F_{genotype} (1,23) = −6.9, $P = 1.00$; F_{location} (1,23) = 9.59, $P = 0.005$; $F_{\text{genotype} \times \text{location}}$ (1,23) = 0.04, $P = 0.83$) (Fig. 3c). We assessed motor learning by measuring improvement in ability to remain on an accelerating rotarod over successive trials (Fig. 3d). There was no genotype difference in improvement of performance over time (F_{genotype} (1,84) = 0.002, $P = 0.96$; F_{session} (4,84) = 33.29, $P < 0.0001$; $F_{\text{genotype} \times \text{session}}$ (4,84) = 1.53, $P = 0.20$), indicating that *Prkcz*$^{-/-}$ and wild-type mice learned this task equally well. We also assessed drug reward memory in male mice by measuring cocaine-conditioned place preference (CPP)[15]. Wild-type ($t = 3.838$, $P = 0.006$) and *Prkcz*$^{-/-}$ mice ($t = 3.645$, $P = 0.01$) spent significantly more time in the cocaine-paired chamber after conditioning and there was no genotype difference in the cocaine CPP index ($t = 0.153$, $P = 0.88$) (Fig. 3e). Because ZIP injection into the nucleus accumbens can erase cocaine reward memory in rats[3], we tested whether ZIP reduces cocaine reward memory in male and female *Prkcz*$^{-/-}$ mice using the same cocaine treatment protocol[3]. We found that compared with saline, ZIP impaired cocaine CPP in *Prkcz*$^{-/-}$ mice ($t = 2.258$, $P = 0.04$) (Fig. 3f).

The finding that ZIP inhibits memory in *Prkcz*$^{-/-}$ mice suggests that its effect on memory maintenance[1–9] occurs through PKM-ζ-independent mechanisms. Recently, the specificity of ZIP and chelerythrine for inhibiting PKM-ζ has been called into question[16–19]. Part of

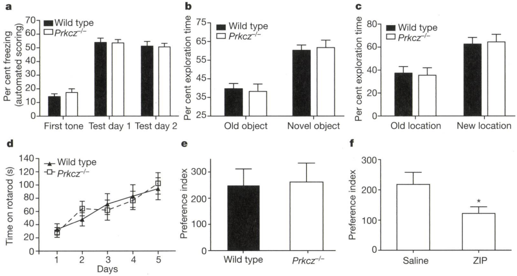


Figure 3 | Intact learning and memory in *Prkcz*$^{-/-}$ mice. **a**, In cued fear conditioning, wild-type ($n = 20$) and *Prkcz*$^{-/-}$ mice ($n = 27$) showed similar levels of freezing during all three sessions. **b**, During the novel object task, *Prkcz*$^{-/-}$ ($n = 17$) and wild-type mice ($n = 16$) spent more time exploring the new object compared with the old object. There was no genotype difference in exploration of old or novel object. **c**, In the spatial memory task, *Prkcz*$^{-/-}$ ($n = 13$) and wild-type mice ($n = 12$) spent more time exploring the new location compared with the old location and there was no genotype difference in exploration of either location. **d**, *Prkcz*$^{-/-}$ ($n = 11$) and wild-type mice ($n = 12$) remained on the accelerating rotarod for similar amounts of time and showed similar improvement in this task over successive trials. **e**, In the CPP test, *Prkcz*$^{-/-}$ ($n = 7$) and wild-type mice ($n = 8$) showed similar preference for the cocaine-paired chamber when tested one day after the last conditioning session. **f**, Compared with injection of saline ($n = 9$), injection of ZIP ($n = 13$) into the nucleus accumbens significantly reduced expression of cocaine CPP in *Prkcz*$^{-/-}$ mice. All data are shown as mean ± s.e.m. $^*P = 0.04$ by two-tailed *t*-test.

this concern arises because the PKC-ζ and -ι pseudosubstrate peptide sequences are identical (SIYRRGARRWRKL), and PKC-ι is widely expressed in the nervous system[20]. To determine the specificity of ZIP, scrambled ZIP and chelerythrine for PKM-ζ, we tested these compounds in an *in vitro* kinase assay using purified PKM-ζ, PKC-ζ and PKC-ι (Table 1 and Supplementary Fig. 3). We found that both ZIP and scrambled ZIP inhibited PKM-ζ in the micromolar range, which is less potent than recently reported[21]. There was only a 7.3-fold difference in $K_{\rm i}$ values between ZIP and scrambled ZIP, and this modest difference in $K_{\rm i}$ values suggests that scrambled ZIP is not an ideal control peptide for ZIP inhibition of PKM-ζ. Interestingly, ZIP and scrambled ZIP were equally potent inhibitors of PKC-ι and PKC-ζ compared with PKM-ζ (Table 1). Chelerythrine was a weak inhibitor of PKM-ζ when assayed without dithiothreitol (DTT), and lost all inhibitory activity when 1 mM DTT was included (Table 1 and Supplementary Fig. 3d). These results question the use of ZIP and chelerythrine as specific inhibitors of PKM-ζ.

Our *in vitro* studies indicate that the current pharmacological reagents commonly used to inhibit PKM-ζ are not specific for this kinase. More importantly, our *in vivo* studies with *Prkcz*$^{-/-}$ mice indicate that PKM-ζ is not required for long-term memory and that ZIP can impair memory through mechanisms that do not involve PKM-ζ. These findings cast doubt on the importance of PKM-ζ in the maintenance of long-term memory.

Table 1 | *In vitro* inhibitory activity of compounds against purified atypical PKC isozymes

Compound	PKC-ζ (μM)	PKM-ζ (μM)	PKC-ι (μM)
ZIP	1.70 (1.14–2.54)	2.11 (1.91–2.33)	1.43 (1.21–1.68)
Scrambled ZIP	5.51 (3.12–9.71)	15.4 (14.7–16.2)	4.92 (3.48–6.96)
Chelerythrine	NA	75.97 (68.44–84.25)	NA

Values are $K_{\rm i}$, 95% CI. NA, not assessed.

METHODS SUMMARY

Generation and testing of *Prkcz*$^{-/-}$ mice. A targeting construct containing a 1.1-kilobase (kb) floxed region of exon 9 was used to generate chimaeric mice that were crossed to produce wild-type and *Prkcz*$^{-/-}$ littermates. To detect PKC-ζ and PKM-ζ, an anti-PKC-ζ antibody (T. Sacktor) was used. Anti-phospho-PKC (pan) (ζThr 410), which detects PKC-ι phosphorylated at T411, was purchased from Cell Signaling Technology, as were antibodies to detect PKC-α, PKC-δ, PKC-ι and PKC-θ. Anti-PKC-β, PKC-γ and PKC-η antibodies were purchased from BD Transduction Laboratories. Anti-PKC-ε antibody was previously generated (SN134)[22].

Behaviour. The behavioural screen was based on methods described previously[13]. Methods for other behavioural tests are available in the online version of the manuscript. Data were examined for normality using a D'Agostino and Pearson omnibus normality test. Light–dark box and elevated plus maze results were analysed by two-factor analysis of variance (ANOVA) with a Bonferroni post-hoc test. The relationship between machine and hand scoring of fear conditioning was analysed by calculating a Pearson product-moment correlation coefficient. In all tests of learning and memory we did not detect a sex difference; to increase power and the possibility of detecting a genotype difference, we combined results from male and female mice and analysed these data by ANOVA, *t*-, or Mann–Whitney *U*-tests.

Full Methods and any associated references are available in the online version of the paper.

Received 18 August; accepted 20 November 2012.

Published online 2 January 2013.

1. Hardt, O., Migues, P. V., Hastings, M., Wong, J. & Nader, K. PKMζ maintains 1-day- and 6-day-old long-term object location but not object identity memory in dorsal hippocampus. *Hippocampus* **20,** 691–695 (2010).
2. He, Y. Y. *et al.* PKMζ maintains drug reward and aversion memory in the basolateral amygdala and extinction memory in the infralimbic cortex. *Neuropsychopharmacology* **36,** 1972–1981 (2011).
3. Li, Y. Q. *et al.* Inhibition of PKMζ in nucleus accumbens core abolishes long-term drug reward memory. *J. Neurosci.* **31,** 5436–5446 (2011).
4. Migues, P. V. *et al.* PKMζ maintains memories by regulating GluR2-dependent AMPA receptor trafficking. *Nature Neurosci.* **13,** 630–634 (2010).
5. Parsons, R. G. & Davis, M. Temporary disruption of fear-potentiated startle following PKMz inhibition in the amygdala. *Nature Neurosci.* **14,** 295–296 (2011).
6. Pastalkova, E. *et al.* Storage of spatial information by the maintenance mechanism of LTP. *Science* **313,** 1141–1144 (2006).

7. Serrano, P. *et al.* PKMζ maintains spatial, instrumental, and classically conditioned long-term memories. *PLoS Biol.* **6,** e318 (2008).
8. Shabashov, D., Shohami, E. & Yaka, R. Inactivation of PKMζ in the NAc shell abolished cocaine-conditioned reward. *J. Mol. Neurosci.* **47,** 546–553 (2011).
9. Shema, R., Sacktor, T. C. & Dudai, Y. Rapid erasure of long-term memory associations in the cortex by an inhibitor of PKM zeta. *Science* **317,** 951–953 (2007).
10. Hernandez, A. I. *et al.* Protein kinase M ζ synthesis from a brain mRNA encoding an independent protein kinase C ζ catalytic domain. Implications for the molecular mechanism of memory. *J. Biol. Chem.* **278,** 40305–40316 (2003).
11. Chou, M. M. *et al.* Regulation of protein kinase C ζ by PI 3-kinase and PDK-1. *Curr. Biol.* **8,** 1069–1078 (1998).
12. Le Good, J. A. *et al.* Protein kinase C isotypes controlled by phosphoinositide 3-kinase through the protein kinase PDK1. *Science* **281,** 2042–2045 (1998).
13. Crawley, J. N. Behavioral phenotyping strategies for mutant mice. *Neuron* **57,** 809–818 (2008).
14. Clarke, J. R., Cammarota, M., Gruart, A., Izquierdo, I. & Delgado-Garcia, J. M. Plastic modifications induced by object recognition memory processing. *Proc. Natl Acad. Sci. USA* **107,** 2652–2657 (2010).
15. Brabant, C., Quertemont, E. & Tirelli, E. Influence of the dose and the number of drug-context pairings on the magnitude and the long-lasting retention of cocaine-induced conditioned place preference in C57BL/6J mice. *Psychopharmacology (Berl.)* **180,** 33–40 (2005).
16. Lisman, J. Memory erasure by very high concentrations of ZIP may not be due to PKM-ζ. *Hippocampus* **22,** 648–649 (2011).
17. Davies, S. P., Reddy, H., Caivano, M. & Cohen, P. Specificity and mechanism of action of some commonly used protein kinase inhibitors. *Biochem. J.* **351,** 95–105 (2000).
18. Wu-Zhang, A. X., Schramm, C. L., Nabavi, S., Malinow, R. & Newton, A. C. Cellular pharmacology of protein kinase Mζ (PKMζ) contrasts with its *in vitro* profile: implications for PKMζ as a mediator of memory. *J. Biol. Chem.* **287,** 12879–12885 (2012).
19. Nishikawa, K., Toker, A., Johannes, F. J., Songyang, Z. & Cantley, L. C. Determination of the specific substrate sequence motifs of protein kinase C isozymes. *J. Biol. Chem.* **272,** 952–960 (1997).
20. Naik, M. U. *et al.* Distribution of protein kinase Mζ and the complete protein kinase C isoform family in rat brain. *J. Comp. Neurol.* **426,** 243–258 (2000).
21. Yao, Y. *et al.* Matching biochemical and functional efficacies confirm ZIP as a potent competitive inhibitor of PKMζ in neurons. *Neuropharmacology* **64,** 37–44 (2013).
22. Choi, D. S., Wang, D., Dadgar, J., Chang, W. S. & Messing, R. O. Conditional rescue of protein kinase C ε regulates ethanol preference and hypnotic sensitivity in adult mice. *J. Neurosci.* **22,** 9905–9911 (2002).

Supplementary Information is available in the online version of the paper.

Acknowledgements We thank V. N. Kharazia and A. J. Lean for assistance. This work was supported by National Institutes of Health grant AA017072 (R.O.M.), a Canadian Institute of Health Research Post-doctoral Fellowship (A.M.L.), and funds provided by the State of California for medical research on alcohol and drug abuse to UCSF.

Author Contributions A.M.L. designed experiments, collected and analysed data, and wrote the manuscript. B.R.K., J.P.L., M.E.Z., C.Q., T.M. and S.C.F.-W. collected and analysed data. D.W. collected and analysed the data from the *in vitro* kinase assays. J.D. produced the constructs for the kinase assays and for generation of the mutant mice, and genotyped the mice. R.O.M. designed experiments, analysed data and co-authored the manuscript.

Author Information Reprints and permissions information is available at www.nature.com/reprints. The authors declare no competing financial interests. Readers are welcome to comment on the online version of the paper. Correspondence and requests for materials should be addressed to R.O.M. (romes@gallo.ucsf.edu or romessing@austin.utexas.edu).

LETTER

doi:10.1038/nature11802

PKM-ζ is not required for hippocampal synaptic plasticity, learning and memory

Lenora J. Volk[1]*, Julia L. Bachman[1]*, Richard Johnson[1], Yilin Yu[1] & Richard L. Huganir[1]

Long-term potentiation (LTP), a well-characterized form of synaptic plasticity, has long been postulated as a cellular correlate of learning and memory. Although LTP can persist for long periods of time[1], the mechanisms underlying LTP maintenance, in the midst of ongoing protein turnover and synaptic activity, remain elusive. Sustained activation of the brain-specific protein kinase C (PKC) isoform protein kinase M-ζ (PKM-ζ) has been reported to be necessary for both LTP maintenance and long-term memory[2]. Inhibiting PKM-ζ activity using a synthetic zeta inhibitory peptide (ZIP) based on the PKC-ζ pseudosubstrate sequence reverses established LTP *in vitro* and *in vivo*[3,4]. More notably, infusion of ZIP eliminates memories for a growing list of experience-dependent behaviours, including active place avoidance[4], conditioned taste aversion[5], fear conditioning and spatial learning[6]. However, most of the evidence supporting a role for PKM-ζ in LTP and memory relies heavily on pharmacological inhibition of PKM-ζ by ZIP. To further investigate the involvement of PKM-ζ in the maintenance of LTP and memory, we generated transgenic mice lacking PKC-ζ and PKM-ζ. We find that both conventional and conditional PKC-ζ/PKM-ζ knockout mice show normal synaptic transmission and LTP at Schaffer collateral–CA1 synapses, and have no deficits in several hippocampal-dependent learning and memory tasks. Notably, ZIP still reverses LTP in PKC-ζ/PKM-ζ knockout mice, indicating that the effects of ZIP are independent of PKM-ζ.

PKM-ζ is transcribed from an alternate internal promoter within the *Prkcz* gene encoding only the catalytic domain, rendering it free from pseudosubstrate auto-inhibition and PKC-ζ regulatory signal transduction pathways[7,8]. To generate mice lacking PKM-ζ we targeted exon 11 of the PKC-ζ catalytic domain for excision by Cre recombinase (Supplementary Fig. 1a). Crossing floxed PKC-ζ alleles (PKC-$\zeta^{\mathrm{fl/fl}}$) to cytomegalovirus (CMV)-Cre transgenic mice resulted in deletion of both PKC-ζ and PKM-ζ, as verified by southern and western blot analysis (Fig. 1b, c). PKC-ζ/PKM-ζ KO mice were both viable and fertile, and there were no anatomical abnormalities in the brains of PKC-ζ/PKM-ζ KO mice (Fig. 1a). Protein expression levels of a number of synaptic proteins as well as different PKC isozymes, including the other atypical isoform PKC-λ/ι, were unchanged in PKC-ζ/PKM-ζ KO mice (Fig. 1c and Supplementary Fig. 1b–d). Analysis of basal synaptic transmission revealed no difference in input/output relationships in adult PKC-ζ/PKM-ζ KO mice compared to wild-type littermates (Supplementary Fig. 2a, b). Paired-pulse facilitation (PPF) was also unaffected, indicating normal presynaptic release probability in PKC-ζ/PKM-ζ KO mice (Supplementary Fig. 2c). Surprisingly, we found theta-burst stimulation (TBS) resulted in normal induction and maintenance of LTP in PKC-ζ/PKM-ζ KO mice, with a magnitude and time course similar to wild-type littermates (Fig. 1d). Some evidence suggests the dependence of LTP maintenance on PKM-ζ activity can vary with the induction protocol[9]. Therefore, we also tested high frequency stimulation (HFS) and again found no LTP deficit in PKC-ζ/PKM-ζ KO mice (Fig. 1e). As PKM-ζ is proposed to be necessary particularly for the late, protein-synthesis-dependent phase of LTP[2], we used the protein-synthesis inhibitor emetine to verify that both stimulation protocols produce forms of LTP that are protein-synthesis-dependent in wild-type mice (Supplementary Fig. 3a, b). Furthermore, TBS-LTP in PKC-ζ/PKM-ζ KO mice is also protein-synthesis-dependent on a similar time scale as observed in wild-type mice (Supplementary Fig. 3c).

To limit possible compensatory effects resulting from PKC-ζ/PKM-ζ deletion during development, we also created conditional PKM-ζ knockout mice by crossing PKC-$\zeta^{\mathrm{fl/fl}}$ to CaMKII-CreER$^{\mathrm{T2}}$ mice. These mice express Cre recombinase which becomes active only upon addition of tamoxifen driven by the *CamkIIα* (also known as *Camk2a*) promoter[10]. This inducible system allows us to delete PKM-ζ specifically in forebrain neurons of adult animals at will. Adult animals (8–10 weeks old) were injected with tamoxifen (2 mg per day for 5 days) and analysed 2 to 3 weeks later. We observed an ~80% reduction in PKM-ζ levels in the dorsal hippocampus of PKC-$\zeta^{\mathrm{fl/fl}}$; CaMKII-CreER$^{\mathrm{T2}}$ mice (Fig. 2a) compared to controls (PKM-$\zeta^{\mathrm{fl/fl}}$ treated with tamoxifen or PKC-$\zeta^{\mathrm{fl/fl}}$; CaMKII-CreER$^{\mathrm{T2}}$ treated with vehicle). Remaining PKM-ζ expression is likely limited to interneurons, as evidence suggests PKM-ζ is not expressed in glia[8] (and data not shown). In acute hippocampal slices from conditional PKM-ζ KO mice we saw no differences in input/output curves or PPF (Supplementary Fig. 2d–f), indicating that basal synaptic transmission is unaltered in these animals. Consistent with our findings in conventional PKC-ζ/PKM-ζ KO mice, we found no deficits in either TBS- or HFS-LTP in conditional PKM-ζ KO mice (Fig. 2b, c). As we saw no significant differences between the conventional and conditional knockout animals, we used conventional PKC-ζ/PKM-ζ KO mice for the remainder of our experiments.

Recent studies have raised concerns regarding the specificity and appropriate use of ZIP[11–14]. We find that when added after LTP induction, ZIP was indeed able to reverse LTP in wild-type mice, confirming previously published results[3,15]. In our hands, however, not only were tetanized responses reduced below baseline (see also ref. 16), but ZIP also decreased synaptic transmission in non-tetanized control slices (Fig. 3b). The effects of ZIP were concentration-dependent, as neither LTP maintenance nor baseline transmission was reduced by 1 μM ZIP (Supplementary Fig. 4a–c) and 2–2.5 μM ZIP was effective in roughly half of our experiments (Supplementary Fig. 4d, e), while 5 μM ZIP reliably reduced both LTP and baseline transmission (Fig. 3). The reduction in field excitatory postsynaptic potential (fEPSP) amplitude was not due to changes in pre-synaptic release, as we did not observe changes in PPF following ZIP application (data not shown). Furthermore, whereas a scrambled version of ZIP (Scr-ZIP) is commonly used as a control peptide, some evidence suggests it may mimic the effects of ZIP *in vivo*[17]. To our surprise, Scr-ZIP was equally effective at reversing LTP in acute hippocampal slices (Supplementary Fig. 4f, g). Supporting these results, both peptides inhibit purified PKM-ζ and PKC-ζ activity *in vitro*, showing overlapping inhibition curves (Supplementary Fig. 4h, i). These effects of ZIP and scrambled ZIP were seen with peptides from several different vendors as well as with

---

[1] Department of Neuroscience, Howard Hughes Medical Institute, Johns Hopkins University School of Medicine, Baltimore, Maryland 21205, USA.

*These authors contributed equally to this work.

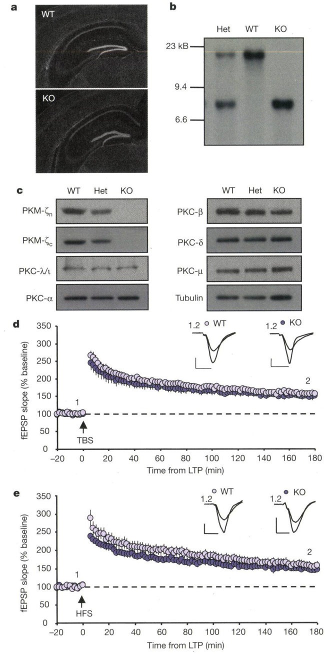


Figure 1 | Normal LTP, gross brain morphology, and PKC isoform expression in conventional PKC-ζ/PKM-ζ KO mice. a, Hippocampal regions stained with DAPI. **b**, Southern blot analysis of wild-type, heterozygous (Het) and homozygous (KO) PKC-ζ/PKM-ζ mice. **c**, Protein expression of PKC isoforms from whole-brain tissue. **d, e**, TBS-LTP (**d**; WT, $n = 11$, 156 ± 6% at 175–180 min; KO, $n = 9$, 149 ± 8%; $P > 0.5$) and HFS-LTP (**e**; WT, $n = 14$, 155 ± 11% at 175–180 min; KO, $n = 13$, 148 ± 6%; $P > 0.5$) are intact and maintained for 3 h in mice lacking PKM-ζ. Data represent mean ± s.e.m. Scale bars, 0.5 mV (vertical), 5 ms (horizontal).

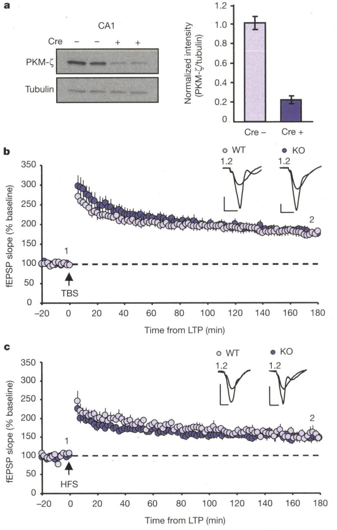


Figure 2 | LTP is intact in conditional PKC-ζ/PKM-ζ knockout mice. a, Western blot (left) and quantification (right) of PKM-ζ protein reduction in conditional KO mice normalized to tubulin (WT, $n = 17$, 101 ± 6.5%; KO, $n = 20$, 22 ± 3.6%). **b, c**, TBS-LTP (**b**; WT, $n = 16$, 180 ± 10% at 175–180 min; KO, $n = 16$, 177 ± 9%; $P > 0.8$) and HFS-LTP (**c**; WT, $n = 6$, 154 ± 11% at 175–180 min; KO, $n = 9$, 148 ± 10%; $P > 0.7$) are intact and maintained for 3 h in mice with PKM-ζ conditionally deleted in adulthood. Data represent mean ± s.e.m. Scale bars, 0.5 mV (vertical), 5 ms (horizontal).

custom synthesized peptides. Importantly, the concentration of ZIP and Scr-ZIP used in our experiments is consistent with that used in previous studies (1–5 μM)[13,15,18].

ZIP and Scr-ZIP use an N-terminal myristoylation group to permeate cell membranes. Several studies in non-neuronal cells have demonstrated that the myristoylation group of ZIP can activate intracellular signalling cascades independent of PKC-ζ activity[19,20]. To control for possible non-specific effects of the myristoylation group, we tested whether myr-PKI, a myristoylated peptide inhibitor of protein kinase A (PKA), would give similar results. Whereas inhibition of PKA can prevent LTP induction, it has no effect on LTP that is already established[21,22]. Application of myr-PKI 1 h after LTP induction had no effect on LTP maintenance or baseline transmission (Fig. 3d). These data indicate that the reduction of LTP by ZIP is not due to the myristoyl group.

ZIP is reported to be highly specific for PKC-ζ/PKM-ζ[3]. It is therefore expected to have no effect on LTP in PKC-ζ/PKM-ζ KO mice. However, not only did ZIP reverse both TBS- and HFS-LTP in PKC-ζ/PKM-ζ KO mice (Fig. 3a, c), but the efficacy of ZIP in PKC-ζ/PKM-ζ KO mice was indistinguishable from that in wild-type (Fig. 3a–c). This raises the question of whether the other atypical PKC isoform, PKC-λ/ι, which contains the same pseudosubstrate sequence, compensates for the lack of PKM-ζ. To date, there is no evidence that PKC-λ/ι produces a truncated, constitutively active form *in vivo*. To test whether alternate products are generated in the absence of PKM-ζ under LTP conditions, we probed tetanized slices from PKC-ζ/PKM-ζ KO mice with a PKC-λ/ι-specific antibody. We found no evidence for the formation of a truncated product 2 h after LTP induction in PKC-ζ/PKM-ζ KO mice (Supplementary Fig. 5a, b). Expression of full-length PKC-λ/ι was also

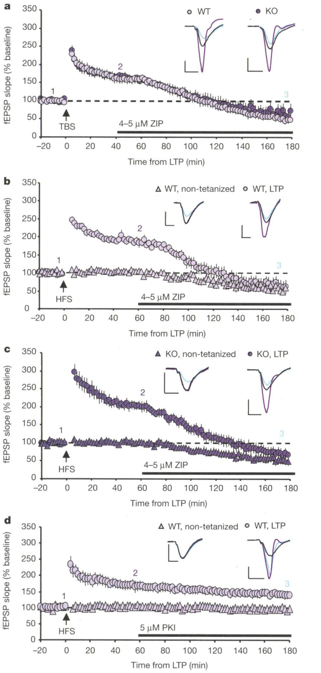


Figure 3 | ZIP is not specific for PKM-ζ. **a**, ZIP is equally effective at reversing established TBS-LTP in WT mice and mice lacking PKM-ζ. Before ZIP application (40 min post-LTP): WT, $n = 8$, 166 ± 12%; KO, $n = 6$, 167 ± 10%, $P > 0.9$. 140 min after ZIP application: WT, 48 ± 7%; KO, 60 ± 18%, $P > 0.5$. **b**, **c**, ZIP decreases both tetanized (HFS) and non-tetanized synaptic responses in WT mice (**b**) and mice lacking PKM-ζ (**c**). Before ZIP application (60 min post-LTP): WT, $n = 8$, tetanized = 181 ± 8%, non-tetanized = 102 ± 4%; KO, $n = 8$, tetanized = 203 ± 12%, non-tetanized = 100 ± 4%, $P > 0.15$ WT vs KO tetanized or non-tetanized. 120 min after ZIP application: WT, tetanized = 60 ± 18%, non-tetanized = 50 ± 9%; KO, tetanized = 67 ± 14%, non-tetanized = 52 ± 11%, $P > 0.6$ WT vs KO tetanized or non-tetanized. **d**, Myristoylated PKI peptide does not affect LTP or basal transmission (tetanized $n = 5$, non-tetanized $n = 4$). Data represent mean ± s.e.m. Scale bars, 0.5 mV (vertical), 5 ms (horizontal).

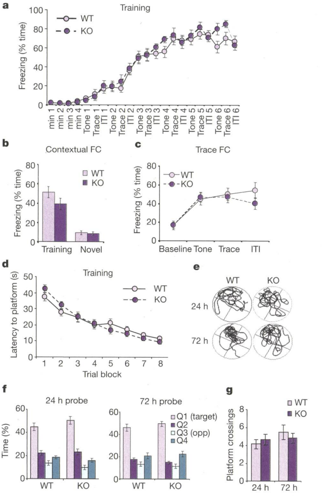


Figure 4 | Hippocampal-dependent learning and memory are intact in PKC-ζ/PKM-ζ KO mice. **a**, Trace fear conditioning elicits freezing in PKC-ζ/PKM-ζ KO ($n = 13$) and WT ($n = 15$) mice with no significant main effect of genotype ($F_{(1,546)} = 0.50, P > 0.48$). ITI, inter-trial interval. **b**, WT and PKC-ζ/PKM-ζ KO mice exhibit similar contextual freezing behaviour (24 h post-training, WT = 51.5 ± 5.7%, KO = 39.3 ± 5.7%, $P_{\mathrm{WT\ vs\ KO}} > 0.16$; significant elevation in post-training vs pre-training freezing, $P < 0.05$ for WT and KO), but show little freezing in a novel context (48 h post-training, WT = 9.3 ± 1.8%, KO = 8.5 ± 1.9%, WT and KO $P > 0.05$ vs pre-training freezing). **c**, Trace fear conditioning is unaffected in PKC-ζ/PKM-ζ KO mice 48 h after training (no significant main effect of genotype, $F_{(1,78)} = 0.33$, $P > 0.57$). **d**, Mean escape latencies during Morris water maze training (WT $n = 21$, KO $n = 17$, no significant main effect of genotype, $F_{(1,252)} = 0.00$, $P > 0.98$). **e**, Representative swim paths during probe trials for WT and PKC-ζ/PKM-ζ KO mice. **f**, Percentage of time spent in each quadrant during probe trials. Both genotypes showed a significant preference for the target quadrant at 24 (WT $n = 21$, KO $n = 17$) and 72 h (WT $n = 17$, KO $n = 15$); $P < 0.001$ (target quadrant vs Q2, Q3 or Q4). There was no significant main effect of genotype at 24 ($F_{(1,108)} = 0.79, P > 0.38$) or 72 h ($F_{(1,90)} = 0.26, P > 0.61$). **g**, Number of platform crossings during probe trials (24 h: WT = 4.14 ± 0.52, KO = 4.64 ± 0.59, $P > 0.5$, 72 h: WT = 5.47 ± 0.8, KO = 4.80 ± 0.54, $P > 0.5$). Data represent mean ± s.e.m.

unchanged. Future work with conditional PKC-λ/ι knockout mice is needed to validate whether PKC-λ/ι has a role in LTP maintenance and could account for the effects of ZIP.

Although the lack of PKM-ζ does not appear to affect canonical forms of LTP in the hippocampus, plasticity mechanisms other than LTP clearly affect learning and memory[23]. Given the plethora of studies supporting the role of PKM-ζ in memory storage, we tested whether deletion of PKM-ζ could affect two different forms of hippocampal-dependent learning and memory. To assess classical associative conditioning, we chose auditory trace fear conditioning[24]. Freezing behaviour of both wild-type and PKC-ζ/PKM-ζ KO mice increased quickly and markedly during the training session (Fig. 4a). Placing the animals back into the training context 24 h later elicited context-evoked freezing behaviour in PKC-ζ/PKM-ζ KO mice that was not significantly different from wild-type littermates (Fig. 4b). On the following day, mice were placed in a novel context and presented with only the tone. Whereas neither wild-type or PKC-ζ/PKM-ζ KO mice exhibited significant freezing behaviour in the novel context, tone presentation resulted in similar levels of elevated freezing in both genotypes (Fig. 4b, c). These data show that PKC-ζ/PKM-ζ KO mice exhibit normal contextual and trace components of aversive conditioning.

To test spatial reference memory in PKC-ζ/PKM-ζ KO mice, we evaluated their performance in the Morris water maze, a task where rodents use spatial cues to locate a hidden platform in a pool of opaque water. Latency to find the platform gradually decreased over the course of eight training blocks for both genotypes, indicating PKC-ζ/PKM-ζ KO mice learned the task as well as wild-type mice (Fig. 4d). Memory retention tests were conducted 24 and 72 h after training. Both wild-type and PKC-ζ/PKM-ζ KO mice spent a significantly greater amount of time in the target quadrant (Fig. 4e, f). In addition, the number of times animals crossed through the platform location was not significantly different between the two groups (Fig. 4g). Contrary to results using ZIP infusion in rats[6], our results demonstrate that global as well as specific location memory is unaffected by the deletion of PKM-ζ.

In conclusion, to investigate the role of PKM-ζ in the maintenance of LTP and learning and memory we generated constitutive and conditional PKC-ζ/PKM-ζ knockout mice. Using both prolonged and acute genetic deletion we find that PKM-ζ is not required for the induction or maintenance of LTP in the CA1 region of the hippocampus. The inhibitory peptide ZIP still reversed LTP in the PKC-ζ/PKM-ζ knockout, indicating that ZIP either has non-specific effects or its action is independent of PKM-ζ. Further studies will be necessary to determine ZIP's specificity and site of action. Moreover, deletion of PKC-ζ/PKM-ζ had no effect on two forms of hippocampal-dependent learning and memory. Intracellular perfusion of PKM-ζ has been shown to enhance synaptic transmission by increasing the number of synaptic AMPA (alpha-amino-3-hydroxy-5-methyl-isoxazole-4-propionic acid)-type glutamate receptors[3,25]. PKM-ζ overexpression has also been shown to enhance memory for conditioned taste aversion[26], suggesting that under some conditions or in select brain regions overexpression of PKM-ζ may regulate synaptic plasticity and memory. However, our results demonstrate that normal synaptic plasticity, learning, and memory can occur in the absence of PKM-ζ and call into question an essential, general role for PKM-ζ in the maintenance of synaptic plasticity and memory.

METHODS SUMMARY

Electrophysiology. Extracellular recordings were made from transverse hippocampal slices (2.5–3.5-month-old mice) submerged in standard ACSF. TBS, 4 trains of 10 bursts at 5 Hz; 1 burst = 4 stimuli at 100 Hz, 10 s inter-train interval. HFS, 4 trains of 100 Hz (1 s duration), 20 s inter-train interval. Stimuli during HFS were set at 75% maximum response. All drugs were bath-applied at the indicated time.

Behaviour. 3–4 month old mice were used for behaviour experiments. Trace fear conditioning was conducted as previously described[27] with the exception that the tone test was conducted on day 4 (48 h after training). A detailed schematic is depicted in Supplementary Fig. 6. Morris water maze training consisted of a total of 40 trials over 5 days, with 4, 60 s trials per session and 2 sessions per day separated by approximately 2 h. On the fifth day the first trial consisted of a 60-s probe trial in which the platform was removed (24-h probe trial). The remaining 8 training trials were conducted after the 24-h probe trial. A second probe trial was performed 72 h after the first probe trial. All animals were housed in standard conditions and treated in accordance with the Johns Hopkins University Animal Care and Use Committee guidelines.

Full Methods and any associated references are available in the online version of the paper.

Received 16 August; accepted 20 November 2012.

Published online 2 January 2013.

1. Abraham, W. C., Logan, B., Greenwood, J. M. & Dragunow, M. Induction and experience-dependent consolidation of stable long-term potentiation lasting months in the hippocampus. *J. Neurosci.* **22,** 9626–9634 (2002).
2. Sacktor, T. C. PKMζ, LTP maintenance, and the dynamic molecular biology of memory storage. *Prog. Brain Res.* **169,** 27–40 (2008).
3. Ling, D. S. *et al.* Protein kinase Mζ is necessary and sufficient for LTP maintenance. *Nature Neurosci.* **5,** 295–296 (2002).
4. Pastalkova, E. *et al.* Storage of spatial information by the maintenance mechanism of LTP. *Science* **313,** 1141–1144 (2006).
5. Shema, R., Sacktor, T. C. & Dudai, Y. Rapid erasure of long-term memory associations in the cortex by an inhibitor of PKMζ. *Science* **317,** 951–953 (2007).
6. Serrano, P. *et al.* PKMζ maintains spatial, instrumental, and classically conditioned long-term memories. *PLoS Biol.* **6,** e318 (2008).
7. Hernandez, A. I. *et al.* Protein kinase Mζ synthesis from a brain mRNA encoding an independent protein kinase Cζ catalytic domain. Implications for the molecular mechanism of memory. *J. Biol. Chem.* **278,** 40305–40316 (2003).
8. Hirai, T., Niino, Y. S. & Chida, K. PKCζII, a small molecule of protein kinase Cζ, specifically expressed in the mouse brain. *Neurosci. Lett.* **348,** 151–154 (2003).
9. Sajikumar, S. & Korte, M. Metaplasticity governs compartmentalization of synaptic tagging and capture through brain-derived neurotrophic factor (BDNF) and protein kinase Mζ (PKMζ). *Proc. Natl Acad. Sci. USA* **108,** 2551–2556 (2011).
10. Erdmann, G., Schutz, G. & Berger, S. Inducible gene inactivation in neurons of the adult mouse forebrain. *BMC Neurosci.* **8,** 63 (2007).
11. Lisman, J. Memory erasure by very high concentrations of ZIP may not be due to PKM-zeta. *Hippocampus* **22,** 648–649 (2012).
12. Wu-Zhang, A. X., Schramm, C. L., Nabavi, S., Malinow, R. & Newton, A. C. Cellular pharmacology of protein kinase Mζ (PKMζ) contrasts with its *in vitro* profile. *J. Biol. Chem.* **287,** 12879–12885 (2012).
13. Sacktor, T. C. & Fenton, A. A. Appropriate application of ZIP for PKMζ inhibition, LTP reversal, and memory erasure. *Hippocampus* **22,** 645–647 (2012).
14. Yao, Y. *et al.* Matching biochemical and functional efficacies confirm ZIP as a potent competitive inhibitor of PKMζ in neurons. *Neuropharmacol.* **64,** 37–44 (2013).
15. Serrano, P., Yao, Y. & Sacktor, T. C. Persistent phosphorylation by protein kinase Mζ maintains late-phase long-term potentiation. *J. Neurosci.* **25,** 1979–1984 (2005).
16. Sajikumar, S., Navakkode, S., Sacktor, T. C. & Frey, J. U. Synaptic tagging and cross-tagging: the role of protein kinase Mζ in maintaining long-term potentiation but not long-term depression. *J. Neurosci.* **25,** 5750–5756 (2005).
17. Kwapis, J. L., Jarome, T., Lonegran, M. & Helmstetter, F. Protein kinase Mzeta maintains fear memory in the amygdala but not in the hippocampus. *Behav. Neurosci.* **123,** 844–850 (2009).
18. Mei, F., Nagappan, G., Ke, Y., Sacktor, T. C. & Lu, B. BDNF facilitates L-LTP maintenance in the absence of protein synthesis through PKMζ. *PLoS ONE* **6,** e21568 (2011).
19. Krotova, K. *et al.* Peptides modified by myristoylation activate eNOS in endothelial cells through Akt phosphorylation. *Br. J. Pharmacol.* **148,** 732–740 (2006).
20. Lim, S. *et al.* A myristoylated pseudosubstrate peptide of PKC-ζ induces degranulation in HMC-1 cells independently of PKC-ζ activity. *Life Sci.* **82,** 733–740 (2008).
21. Denny, J. B., Polan-Curtain, J., Rodriguez, S., Wayner, M. J. & Armstrong, D. L. Evidence that protein kinase M does not maintain long-term potentiation. *Brain Res.* **534,** 201–208 (1990).
22. Huang, Y. Y. & Kandel, E. R. Recruitment of long-lasting and protein kinase A-dependent long-term potentiation in the CA1 region of hippocampus requires repeated tetanization. *Learn. Mem.* **1,** 74–82 (1994).
23. Kim, S. J. & Linden, D. J. Ubiquitous plasticity and memory storage. *Neuron* **56,** 582–592 (2007).
24. McEchron, M. D., Bouwmeester, H., Tseng, W., Weiss, C. & Disterhoft, J. F. Hippocampectomy disrupts auditory trace fear conditioning and contextual fear conditioning in the rat. *Hippocampus* **8,** 638–646 (1998).
25. Ling, D. S., Benardo, L. S. & Sacktor, T. C. Protein kinase Mζ enhances excitatory synaptic transmission by increasing the number of active postsynaptic AMPA receptors. *Hippocampus* **16,** 443–452 (2006).
26. Shema, R. *et al.* Enhancement of consolidated long-term memory by overexpression of protein kinase Mζ in the neocortex. *Science* **331,** 1207–1210 (2011).
27. Makuch, L. *et al.* Regulation of AMPA receptor function by the human memory-associated gene KIBRA. *Neuron* **71,** 1022–1029 (2011).

Supplementary Information is available in the online version of the paper.

Acknowledgements We thank G. Schütz for providing the CaMKII Cre-ER^{T2} mice, M. Gallagher and D. Smith for advice on behaviour and M. Coulter for technical support. We also thank all members of the Huganir lab for discussion and support. This work was supported by grants from the National Institute of Health (NS36715) and the Howard Hughes Medical Institute (to R.L.H.). L.J.V. is supported by a training grant from the National Institute of Health (T32MH15330).

Author Contributions L.J.V., J.L.B. and R.L.H. designed experiments. L.J.V. and J.L.B performed experiments and analysed data. R.J. designed and generated the knockout animals. Y.Y. assisted in mating and genotyping mouse lines. L.J.V., J.L.B. and R.L.H. wrote the manuscript.

Author Information Reprints and permissions information is available at www.nature.com/reprints. The authors declare competing financial interests: details are available in the online version of the paper. Readers are welcome to comment on the online version of the paper. Correspondence and requests for materials should be addressed to R.L.H. (rhuganir@jhmi.edu).

doi:10.1038/nature11747

Asymmetric neurotransmitter release enables rapid odour lateralization in *Drosophila*

Quentin Gaudry[1], Elizabeth J. Hong[1], Jamey Kain[2], Benjamin L. de Bivort[2,3] & Rachel I. Wilson[1]

In *Drosophila*, most individual olfactory receptor neurons (ORNs) project bilaterally to both sides of the brain[1,2]. Having bilateral rather than unilateral projections may represent a useful redundancy. However, bilateral ORN projections to the brain should also compromise the ability to lateralize odours. Nevertheless, walking or flying *Drosophila* reportedly turn towards the antenna that is more strongly stimulated by odour[3–5]. Here we show that each ORN spike releases approximately 40% more neurotransmitter from the axon branch ipsilateral to the soma than from the contralateral branch. As a result, when an odour activates the antennae asymmetrically, ipsilateral central neurons begin to spike a few milliseconds before contralateral neurons, and at a 30 to 50% higher rate than contralateral neurons. We show that a walking fly can detect a 5% asymmetry in total ORN input to its left and right antennal lobes, and can turn towards the odour in less time than it requires the fly to complete a stride. These results demonstrate that neurotransmitter release properties can be tuned independently at output synapses formed by a single axon onto two target cells with identical functions and morphologies. Our data also show that small differences in spike timing and spike rate can produce reliable differences in olfactory behaviour.

To navigate towards an odour, an insect can compare the signals from their two antennae and turn in the direction of the stronger signal. Confounding this strategy (by crossing and fixing the antennae, causing them to be spatially reversed) impairs olfactory navigation in bees, ants and locusts[6–8]. *Drosophila* resemble other insects in using this strategy[3–5], but they are unlike other insects in that they have mainly bilateral ORN projections. In the *Drosophila* brain, both ipsilateral and contralateral ORN axons synapse onto antennal lobe projection neurons (PNs)[9], and ipsilateral and contralateral synapses have approximately the same strength[10]. A minority of *Drosophila* olfactory glomeruli receive unilateral ORN projections[1,11], raising the question of whether it is the unilateral glomeruli that enable lateralization.

To investigate lateralization behaviour, we built a spherical treadmill to measure olfactory behaviour in walking *Drosophila* (Fig. 1a and Supplementary Fig. 1). To prevent the head from moving, it was glued to the body, and the two antennae were independently stimulated with odour. When we delivered fermented peach volatiles to one antenna and clean air to the other, flies on the treadmill made a fictive turn towards the odour (Fig. 1b, c). When odour or clean air was delivered symmetrically to the two antennae, flies continued to walk straight.

Fruit volatiles typically activate multiple ORN types[12]. Therefore, we next used a monomolecular odour (pentanoic acid) to target a glomerulus that receives bilateral ORN input more specifically. This odour has just one known high-affinity receptor in the antenna[11,12]. This receptor corresponds to glomerulus DM6, which receives bilateral ORN innervation[2,9]. Pentanoic acid elicited turning behaviour as robust as that elicited by peach volatiles (Fig. 1d). Turning was significantly reduced by a mutation (*Orco*[2]), which silences bilateral ORNs, including the DM6 ORNs (Fig. 1d). The ORN types that project unilaterally do not express the *Orco* gene[11,13], and so this result implies that asymmetric input to strictly bilateral glomeruli can produce turning.

To target a single bilateral ORN type more selectively, we used an optogenetic approach (Fig. 1e). We expressed channelrhodopsin-2 (ChR2) in the ORNs that express the olfactory receptor Or42b and project to glomerulus DM1. These ORNs fired a burst of 3 or 4 spikes during a 50-ms light pulse directed at the antenna with a fine fibre-optic filament (Fig. 1f). In the same fly, light had no effect on adjacent ORNs that did not express ChR2 (Fig. 1f). Illuminating one antenna produced a turn towards the stimulus (Fig. 1g and Supplementary Fig. 2). Conversely, light offset produced a compensatory turn in the other direction (Fig. 1g, i and Supplementary Note 1). The turns evoked by ChR2 in DM1 ORNs were as large as the turns evoked by ChR2 in most ORN types (under the control of the *Orco-Gal4* line; Fig. 1h). Thus, bilateral ORNs can support lateralization behaviour, and turning can be elicited by just a few spikes per neuron in one ORN type.

The temporal control permitted by the optogenetic approach enabled us to precisely determine the latency of the behavioural response. Notably, turning began within approximately 70 ms of the ORN response onset (Fig. 1j). This is faster than the fly's stride period[14] (about 100 ms, Supplementary Fig. 1). Thus, there must be a rapid mechanism in the brain for extracting lateralized information from sensory neurons.

We next asked how asymmetric odour stimuli are encoded at the level of PN spikes. We removed one antenna in order to lateralize the odour stimulus, and we made simultaneous cell-attached recordings from PNs ipsilateral and contralateral to the intact antenna (Fig. 2a, b). We recorded from pairs of PNs in the same glomerulus on different sides of the brain ('sister PNs'), using green fluorescent protein (GFP) to target our electrodes to PNs in glomerulus DM6 or DM1. We used odours that preferentially activate either DM6 ORNs or DM1 ORNs (pentanoic acid or dilute ethyl acetate[12]). In these experiments, we found a small but consistent asymmetry in PN odour responses. This was apparent in the latency to the first odour-evoked spike: the first ipsilateral spike occurred 2.47 ± 0.70 ms earlier than the first contralateral spike in DM6 PNs, and 1.01 ± 0.41 ms earlier in DM1 PNs ($n = 16$ and 6; Fig. 2c). The latency difference between ipsilateral and contralateral spikes was statistically significant for DM6 PN pairs, but this difference was not significant for DM1 PN pairs ($P < 0.005$ and $P = 0.06$, respectively; Wilcoxon signed-rank tests). In addition, we found an asymmetry in odour-evoked firing rates: ipsilateral firing rates were on average about 50% higher than contralateral firing rates for DM6 PNs, with a larger asymmetry for DM1 PNs (Fig. 2d–g). The asymmetry was observed even during the spontaneous firing of these cells, and it was proportionately similar for all odour concentrations (Fig. 2d, g). The asymmetry in spikes must be due to an asymmetry in synaptic currents. Indeed, asymmetric stimulation of the antennae produced systematically larger inward currents in ipsilateral versus contralateral PNs (Supplementary Fig. 3).

Synaptic currents in PNs reflect the combined effects of ORNs and local neurons, most of which release GABA (γ-aminobutyric acid)[15–17].

---

[1] Department of Neurobiology, Harvard Medical School, 220 Longwood Avenue, Boston, Massachusetts 02115, USA. [2] Rowland Institute, Harvard University, 100 Edwin Land Boulevard, Cambridge, Massachusetts 02142, USA. [3] Department of Organismic and Evolutionary Biology and Center for Brain Science, Harvard University, 52 Oxford Street, Cambridge, Massachusetts 02138, USA.

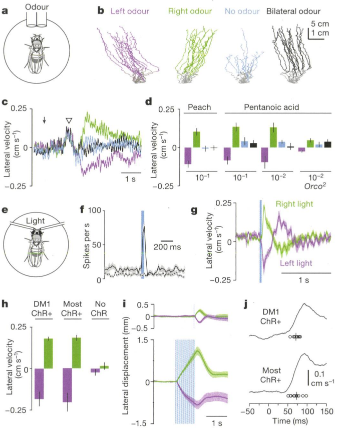


Figure 1 | *Drosophila* can lateralize odours based on bilateral receptor-neuron input to a single pair of glomeruli. a, Schematic diagram of a fly on a spherical treadmill, with odour tubes directed at each antenna. **b**, Representative running trajectories from a single fly. Each trace is a different trial lasting 20 s. The grey portion of each trace indicates the pre-odour baseline period. Flies turn towards lateralized odour (fermented peach extract), but otherwise tend to run straight. **c**, Time course of mean lateral velocity in olfactory stimulation experiments ($n = 9$ flies). Positive values denote rightward turns and negative values denote leftward turns. Arrow indicates the onset of air flow through the tubing. The open arrowhead shows where clean air from the tip of the olfactometer first reaches the flies; this elicits a rightward turn that reflects either a systematic asymmetry in the tethering of the flies or an inherent handedness in the flies. Once this clean air is evacuated, odorized air elicits asymmetrical turning in the fly. Oscillations in lateral velocity are caused by the fly's stride rhythm (Supplementary Fig. 1). Colours as in **b**. **d**, Mean lateral velocity (± s.e.m.) is significantly different for right versus left odour (green versus magenta bars). This was true for peach odour (10^{-1} dilution, $P < 0.005$, $n = 9$ flies, Wilcoxon signed-rank test). It was also true for pentanoic acid (10^{-1} and 10^{-2} dilutions, $n = 8$ and 12, $P < 0.01$ and $P < 0.005$, Wilcoxon signed-rank tests). In *Orco*^{2} mutant flies, responses to right and left odour were still significantly different ($n = 12$ flies, $P < 0.005$, Wilcoxon signed-rank test) but were much smaller. The average difference between left and right (computed within each fly) is significantly smaller in *Orco*^{2} flies compared to control flies ($P < 0.05$, Mann–Whitney *U*-test). The residual response in the mutants is probably due to the ORNs that do not rely on *Orco* (refs 11, 13). **e**, Schematic diagram of a fly on a spherical treadmill, with a fibre-optic light guide directed at each antenna. **f**, Time course of mean light-evoked firing rate in DM1 ORNs (ChR+, solid line) versus other ORN types (ChR−, dashed line). Shaded bands here and elsewhere represent ± s.e.m. **g**, Time course of mean lateral velocity during optogenetic stimulation in flies in which DM1 ORNs are ChR2+ **h**, Mean lateral velocity is significantly different for left and right antennal illumination in flies in which DM1 ORNs are ChR2+, or in flies in which most ORNs are ChR2+ ($n = 10$ and 12, $P < 0.005$ and $P < 5 \times 10^{-4}$, Wilcoxon signed-rank tests), but not in flies in which no ORNs express ChR2 ($P = 0.13$, $n = 10$). To express ChR2 in most ORNs, we used the *Orco-Gal4* line. Flies that lack ChR2 expression have the *UAS-ChR2* transgene but no Gal4 transgene. **i**, Time course of mean lateral displacement in response to a single light pulse (top) or a train of pulses (bottom). After the train, the magnitude of the compensatory turn is larger ($P < 0.005$, *t*-test), but the flies are also significantly less accurate in returning to their original running trajectory ($P < 0.005$, *t*-test). **j**, Time course of absolute mean lateral velocity on an expanded scale around the time of light onset (0 ms). Open circles show the turning latency for each fly. The mean latency across flies is shown as a vertical bar.

We asked whether the asymmetry in PN activity requires GABAergic inhibition, by bath-applying $GABA_A$ and $GABA_B$ receptor antagonists. These antagonists elevated PN firing rates, but the difference between ipsilateral and contralateral PN firing rates was not significantly altered (Fig. 3a, b).

Our results suggest that the asymmetry originates at the level of ORN input to PNs. To compare ORN–PN synapses in ipsilateral versus contralateral PNs, we made simultaneous whole-cell recordings of spontaneous excitatory postsynaptic currents (sEPSCs) in pairs of sister PNs. These sEPSCs are known to arise from ORN–PN synapses[9,10]. Each ORN–PN synapse consists of many release sites with high release probability, and spike-evoked synaptic events are therefore large and reliable[10]. In recordings from sister PNs in glomerulus DM6, almost all sEPSCs occurred in a paired fashion (Fig. 3c). Almost 100% of the sEPSCs in one cell had a corresponding paired event in the other cell, and this percentage was not significantly different in ipsilateral versus contralateral PNs (98.3 ± 0.6% versus 96.4 ± 1.3%, $n = 15$ pairs, $P = 0.11$; Wilcoxon signed-rank test). This result indicates that the number of unitary ORN–PN synaptic connections is essentially identical on the ipsilateral and contralateral sides, and it also indicates that action-potential conduction failures do not occur to any notable degree in the contralateral axon branches of ORNs, although we cannot completely exclude the idea that this occurs at a low rate.

These dual whole-cell recordings enable us to measure the mean difference in ipsilateral versus contralateral sEPSC arrival times (0.80 ± 0.51 ms). The difference in arrival times represents the delay imposed by axonal conduction between sister glomeruli. This delay is less than the delay between the first contralateral odour-evoked spike and the first ipsilateral odour-evoked spike in glomerulus DM6 (2.47 ms, see above). This makes sense, because the difference in first spike times is the result of two delays: the delay imposed by axonal conduction and the delay that results from less net synaptic excitation in the contralateral PN (Supplementary Fig. 3), which means that the contralateral PN requires more integration time to reach spike threshold.

Comparing the amplitude of paired sEPSCs on an event-by-event basis, we found considerable variation in the relative amplitude of ipsilateral versus contralateral EPSCs (Fig. 3d). This is expected, given that synaptic vesicle release is a stochastic process occurring independently at ipsilateral and contralateral synapses. When many sEPSCs were considered together, there was a consistent asymmetry in the amplitude of sEPSCs (Fig. 3d). Averaged across all experiments, ipsilateral sEPSCs were 39% larger in amplitude than their contralateral counterparts (Fig. 3e). Thus, although ORN spikes reliably invade both ipsilateral and contralateral axon branches, each ORN spike typically has a stronger effect on ipsilateral PNs.

In principle, the asymmetry in sEPSC amplitudes could have either a pre- or postsynaptic locus. To determine whether we can observe this asymmetry at the level of presynaptic release sites, we expressed synaptobrevin–GFP (a marker of presynaptic release sites) in DM6 ORNs. We removed one antenna and allowed 3 days for the cut ORN axons to degenerate, leaving only the axons from the intact antenna. We found that total synaptobrevin fluorescence was on average 41 ± 16% higher on the ipsilateral side, and the ipsilateral:contralateral ratio of synaptobrevin fluorescence was significantly greater than 1 ($P < 0.05$, $n = 5$, *t*-test, Fig. 4a). This suggests an asymmetry in the number or size of neurotransmitter release sites. This is consistent with a previous study reporting that a plasma membrane marker

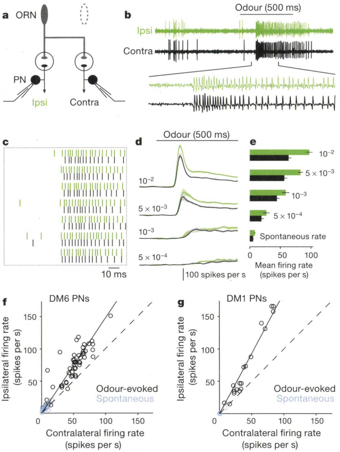


Figure 2 | Lateralized odours produce an asymmetry in spike latency and spike rate in antennal lobe PNs. **a**, Schematic of the set up for simultaneous cell-attached recording. An ORN axon innervates a glomerulus on each side of the midline, where it synapses onto postsynaptic PNs. In these experiments, one antenna was removed to lateralize the odour. **b**, Sample cell-attached recordings from a pair of PNs (postsynaptic to glomerulus DM6). The odour stimulus is a 500-ms pulse of pentanoic acid (10^{-2} dilution). Enlarged segment below is 250 ms. **c**, Raster plot showing the spiking responses of ipsilateral and contralateral DM6 PNs (odour is pentanoic acid at 10^{-2} dilution). Each pair of rows represents a single trial from the same pair of neurons. The raster starts 180 ms after the nominal odour-pulse onset. Note the shorter ipsilateral latency (see text). **d**, Time course of mean firing rates of DM6 PNs to a descending series of pentanoic acid concentrations. **e**, Mean firing rates of DM6 PNs (± s.e.m.) in response to pentanoic acid concentrations. Spontaneous rates are also shown. Mean rate is computed over a 500-ms period starting 100 ms after the olfactometer is activated ($n = 16, 15, 10, 10$ and 21 pairs, in descending order of concentration). **f**, Trial-averaged ipsilateral versus contralateral firing rates for DM6 PNs. Several odour concentrations were used in each experiment, and each point represents a different experiment-concentration combination. Note that the ipsilateral–contralateral difference is present in spontaneous activity (blue symbols). Significance was assessed by fitting a line through the origin to the data for each individual experiment; these slopes were significantly different from unity (mean slope $= 1.47 \pm 0.09$, $P < 10^{-4}$, Mann–Whitney U-test, $n = 21$ experiments). The dashed line is unity; the solid line is the mean of all linear fits in the individual experiments. **g**, Same as **f**, but for DM1 PNs (mean slope $= 1.86 \pm 0.04$, $P < 0.05$, Mann–Whitney U-test, $n = 6$ experiments). Odour stimuli were ethyl acetate at 10^{-12}, 10^{-10} and 10^{-6} dilutions. In DM1 PNs, spontaneous firing rates are close to zero.

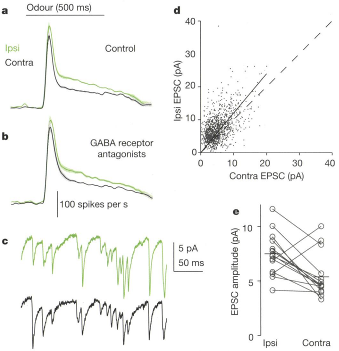


Figure 3 | The asymmetry arises at the level of ORN–PN synapses. **a**, Mean DM6 PN responses to pentanoic acid (10^{-2} dilution) in normal saline. **b**, Same as **a**, but with the addition of the GABA$_A$ receptor antagonist picrotoxin (5 μM) and the GABA$_B$ receptor antagonist CGP54626 (50 μM). Ipsilateral firing rates are significantly higher than contralateral rates ($P < 0.01$), and the antagonists have no significant effect on the ipsilateral–contralateral difference ($P = 0.86$, $n = 10$, two-way analysis of variance (ANOVA) on data from **a** and **b**). The ipsilateral–contralateral difference in peak firing rates becomes somewhat smaller, but because this occurs only at the peak, it is probably due to the near-saturation of PN firing rates. **c**, Whole-cell recordings of sEPSCs from a pair of DM6 PNs. The PN ipsilateral to the intact antenna is in green. **d**, Ipsilateral versus contralateral EPSC amplitudes in a typical pair of DM6 PNs. Each point represents a pair of sEPSCs ($n = 1{,}213$). Dashed line is unity; solid line is a linear fit constrained to intersect the origin. **e**, Group data showing mean sEPSC amplitudes in all pairs of DM6 PNs. Each symbol is a different experiment and horizontal lines represent means across experiments. sEPSC amplitudes are significantly larger in ipsilateral PNs ($n = 15$ pairs, $P < 0.05$, Wilcoxon signed-rank test).

was more abundant on the ipsilateral side[18]. In that study, 1 antenna was severed and imaging was performed 3 days later, as here. Functional remodelling can occur in that time period[18,19], so both findings should be interpreted cautiously.

Next, to visualize presynaptic calcium, we expressed the genetically encoded calcium indicator GCaMP3.0 in ORNs (using the *pebbled-Gal4* line) and used two-photon microscopy to visualize calcium signals in ORN axon terminals. To lateralize the odour, we removed one antenna immediately before the experiment (Fig. 4b). We used pentanoic acid to evoke a fluorescence increase preferentially in glomerulus DM6 (Fig. 4c). We found that the size of the calcium response was significantly greater on the side of the brain ipsilateral to the intact antenna (Fig. 4c, d). This asymmetry did not require feedback from central circuits, because it persisted after washing in mecamylamine (to block nicotinic acetylcholine receptors, which mediate ORN–PN synaptic transmission[10]) along with picrotoxin and CGP54626 hydrochloride (GABA$_A$ and GABA$_B$ receptor antagonists, respectively) (Fig. 4e). In separate experiments, we saw similarly asymmetric calcium signals in glomerulus DL5, using an odour stimulus that is relatively selective for DL5 ORNs[20] (trans-2-hexenal, 10^{-5} dilution, data not shown). We also saw a similar asymmetry in presynaptic currents, using simultaneous bilateral field potential recordings (Supplementary Fig. 4). Together, these results demonstrate that the asymmetry in EPSC amplitudes has a presynaptic origin.

Interestingly, we found a roughly equal asymmetry (~40%) in the amplitude of sEPSCs, the level of synaptobrevin fluorescence, and the odour-evoked GCaMP3 fluorescence change. The simplest explanation for these results is that the ipsilateral arbor is 40% larger than the contralateral arbor. If everything else is equal, then this mechanism

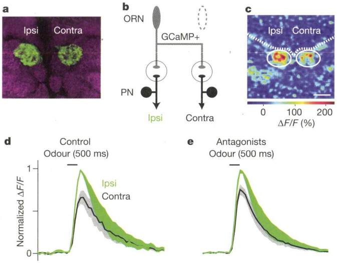


Figure 4 | The asymmetry in ORN–PN synapses has a presynaptic origin. **a**, Synaptobrevin–GFP (green) was expressed in DM6 ORNs to label neurotransmitter release sites, and 1 antenna was removed 3 days before to allow the cut axons to degenerate. The brain is viewed from the anterior face. Magenta shows neuropil contours (nc82 immunofluorescence). Scale bar, 10 μm. **b**, Schematic of calcium-recording configuration. ORNs express GCaMP, and one antenna was removed to lateralize the odour. **c**, Changes in fluorescence in ORN axon terminals in response to pentanoic acid (10^{-2} dilution). Solid circles outline the DM6 glomeruli, and the dashed line shows the anterior boundary of the antennal lobe neuropil. The brain is viewed from the dorsal side. Scale bar, 10 μm. **d**, Mean time course of odour-evoked calcium signals in ORN axons innervating glomerulus DM6. Black bar indicates the timing of the odour stimulus (pentanoic acid at 10^{-2} dilution, 500 ms). **e**, Same as **d**, except after adding picrotoxin, CGP54626, and the nicotinic receptor antagonist mecamylamine (200 μM). Ipsilateral responses are significantly larger than contralateral responses ($P < 0.05$) and the antagonists have no significant effect on the ipsilateral–contralateral difference ($P = 0.75$, $n = 6$, two-way ANOVA on data from panels **d** and **e**).

should produce a proportional change in all these measurements (Supplementary Note 2).

Our results are consistent with the finding that in flies with one antenna removed, 2-deoxyglucose uptake is higher on the side ipsilateral to the intact antenna[21]. A previous study using calcium imaging has reported that asymmetric ORN stimulation can produce stronger signals in ipsilateral PNs than in contralateral PNs[22], which is also broadly consistent with our results. However, in contrast to our results, this asymmetry was restricted to just a few glomeruli. This finding is puzzling, given that a variety of stimuli can produce turning behaviour (Fig. 1 and refs 3–5). The same study also concluded that GABAergic inhibition mediates the ipsilateral–contralateral asymmetry in PNs[22]. These discrepancies may reflect limitations of the imaging techniques used in that study, as well as differences between stimuli or glomeruli.

Another previous study failed to find any significant ipsilateral–contralateral differences in the strength of ORN–PN synapses[10]. However, the measurements in the study involved sequential recordings from sister PNs, rather than simultaneous recordings, and the measurements were of EPSCs arising from a single ORN per PN. Here we have better statistical power because we have made simultaneous recordings from sister PNs, and because we sampled sEPSCs arising from many ORNs. Our finding that ipsilateral and contralateral synapses differ only modestly explains why this difference has been difficult to resolve previously.

It is well-known that a single axon can form neurotransmitter release sites with different properties on different postsynaptic cell types (for example, principal neurons versus interneurons[23]). Here we have shown that ORN axons discriminate between two classes of target cells (left and right) that are morphologically and functionally identical, and that share the same lineage and birth dates[24]. This is reminiscent of the circuitry of the leech midbody ganglion, in which individual mechanoreceptor axons make stronger synapses onto ipsilateral versus contralateral sister cells that are otherwise functionally identical[25]. It is possible that these target cells may be distinguished because right and left sister neurons express different molecular tags[26], thereby allowing a given axon to recognize them as ipsilateral versus contralateral in relation to itself. Alternatively, an axon might form more release sites at proximal locations than at distal locations; this cell-intrinsic mechanism would suffice because contralateral sites are always more distal than ipsilateral sites.

Our results reveal that even small signals in the *Drosophila* nervous system can be behaviourally relevant. The stimuli we used in our optogenetic experiments produced only a slight fractional difference in input to the right and left sides of the brain. This difference amounted to 5% over the 50-ms stimulus period (Supplementary Note 3). The finding that this incremental difference is relevant for behaviour should provide an additional motive for the continued development of sensitive methods for monitoring neural activity in the fly brain.

METHODS SUMMARY

The odour-delivery device used for the olfactory behavioural experiments was specially designed to deliver no lateralized mechanical cues, and we carried out control experiments to confirm that no turning was observed when no odour was present (Supplementary Fig. 5). The spherical treadmill apparatus was constructed by floating a small plastic sphere on a jet of compressed air. We measured the forward velocity (pitch) and lateral velocity (roll) of the sphere in the apparatus by placing the sensor from an optical mouse underneath the sphere. Photostimulation of *Drosophila* ORNs was achieved by butt-coupling a blue light-emitting diode (LED) to a fibre-optic filament (50 μm diameter) and positioning the tip of the fibre approximately 150 μm away from the fly's antenna. In all behavioural experiments using light to stimulate ORNs, the eyes and ocelli of the fly were shielded from light by painting them with ink. *In vivo* extracellular recordings from ORNs and patch-clamp recordings from PNs were carried out as described previously[9,27]. In all electrophysiology or calcium-imaging experiments (except ORN recordings), one antenna was removed just before the experiment by a person who was not the experimenter, and the experimenter remained blind to which side of the brain was ipsilateral to the intact antenna. All analysis was also carried out blind to which side was ipsilateral. Calcium-imaging experiments were performed on a custom-built two-photon microscope. All aggregated data represent mean ± s.e.m. computed across experiments. See Supplementary Methods for details on the spherical treadmill, odour delivery, optogenetic stimuli and analysis of sEPSCs.

Full Methods and any associated references are available in the online version of the paper.

Received 6 May; accepted 7 November 2012.

Published online 23 December 2012.

1. Stocker, R. F., Lienhard, M. C., Borst, A. & Fischbach, K. F. Neuronal architecture of the antennal lobe in *Drosophila melanogaster. Cell Tissue Res.* **262,** 9–34 (1990).
2. Couto, A., Alenius, M. & Dickson, B. J. Molecular, anatomical, and functional organization of the *Drosophila* olfactory system. *Curr. Biol.* **15,** 1535–1547 (2005).
3. Borst, A. & Heisenberg, M. Osmotropotaxis in *Drosophila melanogaster. J. Comp. Physiol. A* **147,** 479–484 (1982).
4. Duistermars, B. J., Chow, D. M. & Frye, M. A. Flies require bilateral sensory input to track odor gradients in flight. *Curr. Biol.* **19,** 1301–1307 (2009).
5. Flugge, C. Geruchliche raumorientierung von *Drosophila melanogaster. Z. Vgl. Physiol.* **20,** 463–500 (1934).
6. Kennedy, J. S. & Moorhouse, J. E. Laboratory observations on locust responses to wind-borne grass odour. *Entomol. Exp. Appl.* **12,** 487–503 (1969).
7. Martin, H. Osmotropotaxis in the honey-bee. *Nature* **208,** 59–63 (1965).
8. Hangartner, W. Spezifität und inaktivierung des spurpheromons von *Lasius fuliginosus* (Latr.) und orientierung der arbeiterinnen im duftfeld. *Z. Vgl. Physiol.* **57,** 103–136 (1967).
9. Kazama, H. & Wilson, R. I. Origins of correlated activity in an olfactory circuit. *Nature Neurosci.* **12,** 1136–1144 (2009).
10. Kazama, H. & Wilson, R. I. Homeostatic matching and nonlinear amplification at genetically-identified central synapses. *Neuron* **58,** 401–413 (2008).
11. Silbering, A. F. *et al.* Complementary function and integrated wiring of the evolutionarily distinct *Drosophila* olfactory subsystems. *J. Neurosci.* **31,** 13357–13375 (2011).

12. Hallem, E. A. & Carlson, J. R. Coding of odors by a receptor repertoire. *Cell* **125,** 143–160 (2006).
13. Benton, R., Vannice, K. S., Gomez-Diaz, C. & Vosshall, L. B. Variant ionotropic glutamate receptors as chemosensory receptors in *Drosophila*. *Cell* **136,** 149–162 (2009).
14. Strauss, R. & Heisenberg, M. Coordination of legs during straight walking and turning in *Drosophila melanogaster*. *J. Comp. Physiol. A* **167,** 403–412 (1990).
15. Chou, Y. H. *et al.* Diversity and wiring variability of olfactory local interneurons in the *Drosophila* antennal lobe. *Nature Neurosci.* **13,** 439–449 (2010).
16. Okada, R., Awasaki, T. & Ito, K. Gamma-aminobuyric acid (GABA)-mediated neural connections in the *Drosophila* antennal lobe. *J. Comp. Neurol.* **514,** 74–91 (2009).
17. Das, A. *et al. Drosophila* olfactory local interneurons and projection neurons derive from a common neuroblast lineage specified by the empty spiracles gene. *Neural Dev.* **3,** 33 (2008).
18. Berdnik, D., Chihara, T., Couto, A. & Luo, L. Wiring stability of the adult *Drosophila* olfactory circuit after lesion. *J. Neurosci.* **26,** 3367–3376 (2006).
19. Kazama, H., Yaksi, E. & Wilson, R. I. Cell death triggers olfactory circuit plasticity via glial signaling in *Drosophila*. *J. Neurosci.* **31,** 7619–7630 (2011).
20. Olsen, S. R., Bhandawat, V. & Wilson, R. I. Divisive normalization in olfactory population codes. *Neuron* **66,** 287–299 (2010).
21. Rodrigues, V. Spatial coding of olfactory information in the antennal lobe of *Drosophila melanogaster*. *Brain Res.* **453,** 299–307 (1988).
22. Agarwal, G. & Isacoff, E. Specializations of a pheromonal glomerulus in the *Drosophila* olfactory system. *J. Neurophysiol.* **105,** 1711–1721 (2011).
23. Pelkey, K. A. & McBain, C. J. Differential regulation at functionally divergent release sites along a common axon. *Curr. Opin. Neurobiol.* **17,** 366–373 (2007).
24. Jefferis, G. S., Marin, E. C., Stocker, R. F. & Luo, L. Target neuron prespecification in the olfactory map of *Drosophila*. *Nature* **414,** 204–208 (2001).
25. Lockery, S. R. & Kristan, W. B. Jr. Distributed processing of sensory information in the leech. II. Identification of interneurons contributing to the local bending reflex. *J. Neurosci.* **10,** 1816–1829 (1990).
26. Chintapalli, V. R. *et al.* Functional correlates of positional and gender-specific renal asymmetry in *Drosophila*. *PLoS ONE* **7,** e32577 (2012).
27. Bhandawat, V., Olsen, S. R., Schlief, M. L., Gouwens, N. W. & Wilson, R. I. Sensory processing in the *Drosophila* antennal lobe increases the reliability and separability of ensemble odor representations. *Nature Neurosci.* **10,** 1474–1482 (2007).

Supplementary Information is available in the online version of the paper.

Acknowledgements We are grateful to M. Dickinson, V. Jayaraman, L. Luo, D. Tracey and L. Vosshall for gifts of fly stocks. A. Baker helped construct and improve the spherical treadmill apparatus. Members of the Wilson laboratory provided feedback on the manuscript. This work was supported by a research project grant from the National Institutes of Health (R01DC008174). R.I.W. is an HHMI Early Career Scientist. B.L.d.B. and J.K. were supported by the Rowland Junior Fellows Program.

Author Contributions Q.G. and R.I.W. designed the experiments. Q.G. carried out all of the experiments, except for the calcium imaging, which was performed by Q.G. and E.J.H., and the synaptobrevin imaging, which was performed by R.I.W. Q.G. analysed the data. J.K. and B.L.d.B. helped to design and build the spherical treadmill apparatus. Q.G. and R.I.W. wrote the manuscript.

Author Information Reprints and permissions information is available at www.nature.com/reprints. The authors declare no competing financial interests. Readers are welcome to comment on the online version of the paper. Correspondence and requests for materials should be addressed to R.I.W. (rachel_wilson@hms.harvard.edu).

doi:10.1038/nature11723

Bacteriophage genes that inactivate the CRISPR/Cas bacterial immune system

Joe Bondy-Denomy[1], April Pawluk[2], Karen L. Maxwell[3] & Alan R. Davidson[1,2]

A widespread system used by bacteria for protection against potentially dangerous foreign DNA molecules consists of the clustered regularly interspaced short palindromic repeats (CRISPR) coupled with *cas* (CRISPR-associated) genes[1]. Similar to RNA interference in eukaryotes[2], these CRISPR/Cas systems use small RNAs for sequence-specific detection and neutralization of invading genomes[3]. Here we describe the first examples of genes that mediate the inhibition of a CRISPR/Cas system. Five distinct 'anti-CRISPR' genes were found in the genomes of bacteriophages infecting *Pseudomonas aeruginosa*. Mutation of the anti-CRISPR gene of a phage rendered it unable to infect bacteria with a functional CRISPR/Cas system, and the addition of the same gene to the genome of a CRISPR/Cas-targeted phage allowed it to evade the CRISPR/Cas system. Phage-encoded anti-CRISPR genes may represent a widespread mechanism for phages to overcome the highly prevalent CRISPR/Cas systems. The existence of anti-CRISPR genes presents new avenues for the elucidation of CRISPR/Cas functional mechanisms and provides new insight into the co-evolution of phages and bacteria.

Predation by phages presents an important challenge to bacterial survival[4], and bacteria have evolved numerous mechanisms to resist phage infection[5]. One such system is the CRISPR/Cas immune system, which is found in 48% of Eubacteria and 95% of Archaea[6]. CRISPR loci contain multiple repeated sequences of approximately 30 base pairs, separated by variable 'spacer' sequences of similar length, which are often identical to segments of phage genomes or other mobile genetic elements[6]. The large single transcript from a CRISPR locus is processed in the repeat regions into small CRISPR RNAs (crRNAs)[7,8] that are complexed with Cas proteins[3,9]. Using the crRNAs as guides, crRNA/Cas complexes cleave foreign DNA molecules at sites bearing complementarity to the crRNAs, thus engendering resistance to phages and other invading DNA molecules[10,11]. Cas proteins also mediate an adaptive function by incorporating short sequences (that is, a new spacer) from newly encountered foreign genomes into CRISPR loci so that these genomes will be destroyed in subsequent encounters[12–14]. The relatively recent elucidation of CRISPR/Cas functions and their obvious similarities to RNA interference (RNAi) systems of eukaryotes have led to vigorous investigation of these systems.

Because phage genes have been discovered that can neutralize most of the prevalent bacterial anti-phage defences[5], the failure to identify genes that counteract the widely occurring CRISPR/Cas systems has been surprising. To search for such 'anti-CRISPR' activity, we investigated the type I–F CRISPR/Cas system[1] of the opportunistic pathogen *Pseudomonas aeruginosa*[14] using a collection of 44 lysogens of *P. aeruginosa* PA14, which each contained a different phage genome (see Methods). In lysogens, phage genomes are integrated into the bacterial genome and are referred to as prophages. Although prophage genes are generally repressed, all prophages have some genes that are actively transcribed. To test whether prophages might express anti-CRISPR activity, we measured the plaquing efficiency of three 'CRISPR-sensitive' phages (JBD18, JBD25 and JBD67) on our collection of lysogens. The CRISPR-sensitive phages fail to replicate on PA14 owing to the action of the CRISPR/Cas system[14], but are able to replicate on PA14 ΔCR/*cas*, which contains no CRISPR loci or *cas* genes (Fig. 1a and Supplementary Fig. 1a). We identified three lysogenic strains, PA14(JBD24), PA14(MP29) and PA14(JBD30), on which the CRISPR-sensitive phages could form plaques very robustly compared with unlysogenized PA14 (Fig. 1a and Supplementary Fig. 1a). Notably, the plaquing efficiency of the CRISPR-sensitive phages on PA14(JBD30) was equivalent to that on the ΔCR/*cas* strain, indicating that the JBD30 prophage caused complete inactivation of the CRISPR/Cas system. The lower plaquing efficiency of the CRISPR-sensitive phages on the other lysogens relative to their plaquing on the ΔCR/*cas* strain may be due to their production of less potent anti-CRISPR activity. However, these prophages also attenuate plaquing through mechanisms independent of the CRISPR/Cas system, as demonstrated by the partial inhibition of plaquing of the control phage, DMS3, which is not affected by the CRISPR/Cas system (Fig. 1a and Supplementary Fig. 1a).

To assess directly the anti-CRISPR activity of the PA14 lysogens, we used a plasmid-based transformation efficiency assay. The sequences within phages that are targeted by the CRISPR/Cas system are called protospacers. To be targeted, a protospacer sequence must be complementary to a specific spacer sequence in the CRISPR locus and possess a correct protospacer adjacent motif[15,16] (Fig. 1b). Protospacer sequences are named according to the spacer sequence that they match in the PA14 genome (Fig. 1b). We constructed plasmids containing targeted protospacer sequences from phages JBD18 (CRISPR2 locus, spacer 1 or CR2_sp1) and JBD25 (CR1_sp1)[14]. The transformation efficiencies of the plasmids bearing protospacers into unlysogenized PA14 were reduced by at least 90% compared with an empty vector control, whereas no difference in transformation efficiency was seen for the three strains containing prophages expressing anti-CRISPR activity, or for the ΔCR/*cas* strain (Fig. 1c). These data confirm that the prophages isolated in our screen inhibit the PA14 CRISPR/Cas system.

Genome comparisons showed that the phages producing anti-CRISPR activity were closely related to each other, and showed high sequence similarity and synteny with the previously characterized *P. aeruginosa* Mu-like phages D3112, DMS3 and MP22 (Supplementary Table 1). An unusual feature of each of these genomes compared with more distantly related Mu-like phages infecting *P. aeruginosa*[17] and other hosts[18] is the presence of diverse atypical genes at a single position in the region encoding phage head proteins (Fig. 1d and Supplementary Fig. 2). We suspected that some of these genes might encode anti-CRISPR activities; thus, 17 of them were cloned and expressed from a high copy number plasmid in PA14 under the control of an arabinose-inducible promoter. Remarkably, expression of eight of these genes (Supplementary Fig. 3) led to marked increases in the plaquing efficiency of the CRISPR-sensitive phages (Fig. 1e and Supplementary Fig. 1b). Each anti-CRISPR gene allowed these three different phages to evade the CRISPR/Cas system even though they bore distinct protospacer targets (Supplementary Table 1), indicating

---

[1]Department of Molecular Genetics, University of Toronto, Toronto, Ontario M5S 1A8, Canada. [2]Department of Biochemistry, University of Toronto, Toronto, Ontario M5S 1A8, Canada. [3]Donnelly Centre for Cellular and Biomolecular Research, University of Toronto, Toronto, Ontario M5S 3E1, Canada.

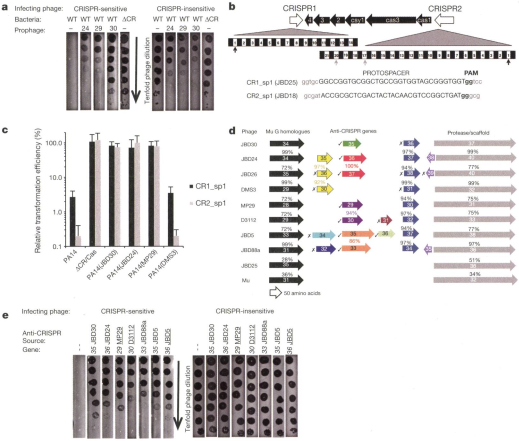

a
Infecting phage: CRISPR-sensitive CRISPR-insensitive
Bacteria: WT WT WT WT ΔCR WT WT WT WT ΔCR
Prophage: – 24 29 30 – – 24 29 30 –
Tenfold phage dilution
b
CRISPR1
CRISPR2
4 3 2 csy1 cas3 cas1
1 2 3 4 5 6 7 8 9 10 11 12 13 14
21 20 19 18 17 16 15 14 13 12 11 10 9 8 7 6 5 4 3 2 1
PROTOSPACER
PAM
CR1_sp1 (JBD25) ggtgcGGCCGGTGCGGCTGCCGGTGGTAGCGGGTGGTggtcc
CR2_sp1 (JBD18) gcgatACCGCGCTCGACTACTACAACGTCCGGCTGATgggcg
c
Relative transformation efficiency (%)
100.0
10.0
1.0
0.1
PA14
ΔCR/Cas
PA14(JBD30)
PA14(JBD24)
PA14(MP29)
PA14(DMS3)
■ CR1_sp1
■ CR2_sp1
d
Phage
Mu G homologues
Anti-CRISPR genes
Protease/scaffold
JBD30
JBD24
JBD26
DMS3
MP29
D3112
JBD5
JBD88a
JBD25
Mu
34 ✓35 ✗36 37
99% 34 35 ✓36 97% 37 38 99% 40
72% 35 97% ✗36 100% ✓37 94% ✗38 ✗39 77% 40
99% 29 92% ✗30 99% ✗31 99% 32
72% 28 ✓29 94% 30 75% 31
72% 29 94% ✓30 ✗31 97% 32 75% 33
72% 33 ✗34 ✓35 ✓36 94% 37 77% 38
99% 31 ✗32 86% ✓33 97% 34 35 97% 36
28% 35 51% 36
36% 31 34% 32
50 amino acids
e
Infecting phage: CRISPR-sensitive CRISPR-insensitive
Anti-CRISPR Source: Gene:
– 35 JBD30 36 JBD24 29 MP29 30 D3112 33 JBD88a 35 JBD5 36 JBD5
– 35 JBD30 36 JBD24 29 MP29 30 D3112 33 JBD88a 35 JBD5 36 JBD5
Tenfold phage dilution

Figure 1 | The CRISPR/Cas system is inhibited by expression of phage genes. **a**, Tenfold dilutions of lysates of a CRISPR-sensitive phage (JBD18) and a CRISPR-insensitive phage (DMS3) were applied to bacterial lawns of wild-type PA14, PA14 lysogens (JBD24, MP29 or JBD30), and PA14 lacking a CRISPR/Cas system (ΔCR). **b**, Schematic of the PA14 CRISPR loci and *cas* gene region is shown. Expanded versions of each CRISPR locus indicate the number of spacers in each, shown with white boxes, each of which is flanked by repeats denoted by black boxes. Black arrows indicate the CRISPR spacers corresponding to protospacers tested in Fig. 1c and grey arrows indicate the CRISPR spacers corresponding to protospacers tested in Supplementary Fig. 11. The DNA sequences of the protospacers tested in Fig. 1c are shown. **c**, Plasmids containing protospacers shown in Fig. 1b were electroporated into the indicated strains. The relative transformation efficiency was calculated by comparison with the transformation efficiency of the cloning vector containing no protospacer insert. Error bars, s.d. of three biological replicates. **d**, The anti-CRISPR genes of the indicated phages are located in the head gene regions of these genomes between genes homologous to the *G* gene of *E. coli* phage Mu (black boxes) and genes encoding the protease/scaffold protein of the phage head (grey boxes; see Supplementary Fig. 2). The percentage identities of the proteins encoded by these genes to representatives from JBD30 are shown. The coloured boxes represent putative anti-CRISPR genes. Boxes of the same colour represent closely related genes and the sequence identity of their encoded proteins is indicated. Genes found to mediate anti-CRISPR activity are indicated by tick ('✓') marks and genes tested but showing no anti-CRISPR activity are marked with a cross ('✗') symbol. Unmarked genes were not tested owing to their high similarity to tested genes. The gene box sizes are proportional to the sizes of the proteins in question (scale bar, 50 amino acids), but the spacing of the genes is not to scale. **e**, The same phages from Fig. 1a were tested on PA14 containing empty vector or plasmids expressing the indicated anti-CRISPR genes. Induction of the plasmid promoter with arabinose was required to produce a maximal effect for some of the anti-CRISPR genes as indicated by underlining. The assays shown in Fig. 1a, e are by necessity taken from different plates. The complete plates for these experiments and experiments with the other CRISPR-sensitive phages are shown in Supplementary Fig. 1.

that the anti-CRISPR genes do not function through protection of specific DNA sequences on the phages. We also found that the anti-CRISPR genes can inhibit the type I–F systems found in other *P. aeruginosa* strains, demonstrating that this phenomenon is not particular to strain PA14 (Supplementary Fig. 4). Finally, we found that the anti-CRISPR genes did not inhibit a type I–E CRISPR/Cas system functioning in *Escherichia coli*[19] (Supplementary Fig. 5). Because type I–E is the most closely related CRISPR/Cas system to type I–F[1], we do not expect that these genes would inhibit the function of any of the other more distantly related CRISPR/Cas systems.

For the type I–F CRISPR/Cas system to function, transcription of pre-crRNA from the CRISPR locus must occur, followed by processing into small crRNAs and incorporation of these RNAs into a complex with Cas proteins[3,7]. The stable maintenance of processed crRNA within the cell requires Cas proteins. Thus, the lack of any *P. aeruginosa* Cas protein except Cas1, which is involved in spacer acquisition, causes a marked reduction in crRNA amounts[20]. As can be seen in Supplementary Fig. 6a, expression of five different anti-CRISPR genes in PA14 caused no change in the amount of processed crRNA molecules as detected by northern blotting. The normal amount of crRNA observed implies that expression of anti-CRISPR genes does not cause a reduction in the expression amounts of the CRISPR loci or *cas* genes. Supporting this finding, we also showed that transcription levels of the *cas* genes *cas3* and *csy3* were unaffected by the anti-CRISPR genes as assessed by reverse transcriptase quantitative PCR (RT–qPCR) (Supplementary Fig. 6b, c). Furthermore, β-galactosidase activity produced from a chromosomally located *csy3::lacZ* fusion gene was not perturbed by expression of any of the anti-CRISPR genes (Supplementary Fig. 6d). We conclude from these experiments that the anti-CRISPR genes exert their effects at a step occurring after formation of the crRNA-Cas complex, and that there is no effect on biogenesis of either the crRNA or Cas proteins.

Despite the common genomic positions of the anti-CRISPR genes in very similar *P. aeruginosa* phages, these eight genes are predicted to encode five different proteins with completely distinct sequences (Supplementary Fig. 3). Sequence similarity searches with each of these proteins yielded fewer than 15 different significant hits in total, of which all but four were proteins encoded in genomes of closely related phages or prophages (Supplementary Table 2). One of the non-phage-associated anti-CRISPR protein homologues, which is 43% identical to the product of gene *33* of phage JBD88a (JBD88a gp33), is encoded in an active pathogenicity island of a highly virulent *P. aeruginosa* clinical isolate that is probably transferred by conjugation between *P. aeruginosa* strains[21]. This island contains four protospacers with correct protospacer adjacent motifs and 100% identity to CRISPR spacers in various *P. aeruginosa* strains[22]. The three other non-phage-associated anti-CRISPR homologues are also found in regions of *Pseudomonas* genomes that may be mobile elements as indicated by presence of genes in these regions encoding homologues of proteins involved in DNA transfer and/or type IV secretion (Supplementary Fig. 7). Thus, these putative bacterial anti-CRISPR genes may increase the fitness for inter-strain transfer of these mobile elements by inactivating the CRISPR/Cas system of a recipient strain.

Because the crRNA/Cas complex is guided by RNA, anti-CRISPR activity might be mediated by a non-coding RNA molecule or a protein encoded by an anti-CRISPR gene. We addressed this issue by performing experiments on JBD30 gene *35*. A nonsense mutation at the third codon and two different frameshift mutations were introduced to the plasmid encoding gene *35*. Each of these mutations abrogated anti-CRISPR activity (Supplementary Fig. 8), implying that translation of this region was required for function. Because these experiments did not rule out a combined role for anti-CRISPR non-coding RNA and protein, two variant genes were synthesized that encoded the same amino-acid sequence as gene *35*, yet had DNA sequences that differed by approximately 35% through variation of codon wobble positions (Supplementary Fig. 9). As shown in Supplementary Fig. 8, each of these synthetic versions of gene *35* imparted full anti-CRISPR activity. These data demonstrate that anti-CRISPR protein is required for anti-CRISPR activity and that a direct mechanistic role for a gene *35*-encoded RNA is unlikely.

The genomes of six of the seven 'anti-CRISPR phages' (that is, those phages bearing active anti-CRISPR genes, Fig. 1d) contain at least one functional protospacer (Supplementary Table 1); thus, their replication should be inhibited by the PA14 CRISPR/Cas system. However, each was able to form plaques on PA14 with 100% efficiency compared with the ΔCR/*cas* strain (Supplementary Fig. 10). Using the transformation efficiency assay, we confirmed that the two protospacers found most commonly in the anti-CRISPR phages were indeed targeted by the PA14 CRISPR/Cas system (Supplementary Fig. 11). These results implied that the anti-CRISPR phages were able to replicate on PA14 because they possess anti-CRISPR genes. To address this hypothesis, a frameshift mutation was introduced into the phage JBD30 anti-CRISPR gene (gene *35*). This mutant phage was unable to replicate on wild-type PA14 but still replicated robustly on the ΔCR/*cas* strain, demonstrating the requirement of the anti-CRISPR gene for replication in cells bearing an intact CRISPR/Cas system (Fig. 2a). To determine whether the introduction of an anti-CRISPR gene into a CRISPR-sensitive phage would allow that phage to evade CRISPR/Cas immunity, we used a CRISPR-sensitive mutant of phage DMS3, called DMS3*m*[14]. This phage possesses a functional protospacer and is very similar in sequence to the anti-CRISPR phages, yet it contains no functional anti-CRISPR gene (Fig. 1c). Taking advantage of the high DNA sequence identity between phages DMS3*m* and JBD30, *in vivo* homologous recombination was used to create a version of DMS3*m* bearing JBD30 gene *35* (Fig. 2b, Supplementary Fig. 12 and Methods). As shown in Fig. 2c, the introduction of this gene into DMS3*m* resulted in a 10^6-fold increase in plaquing efficiency on PA14, clearly demonstrating that an anti-CRISPR gene present on an infecting phage allows that phage to overcome the CRISPR/Cas system. By testing lysogens of the DMS3*m* and JBD30 mutant prophages we found that a protospacer-bearing plasmid transformed cells efficiently only those lysogens in which an intact anti-CRISPR gene was present (Fig. 2d). These assays demonstrate the necessity and sufficiency of the anti-CRISPR gene for inhibition of the CRISPR/Cas system.

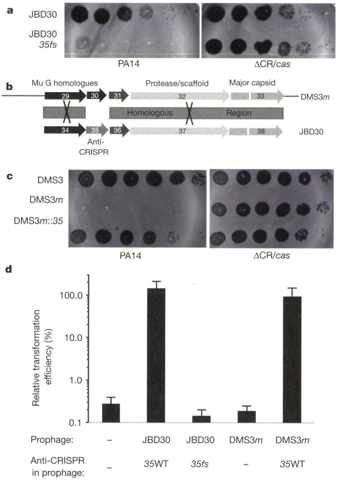


Figure 2 | An anti-CRISPR gene protects phages from the CRISPR/Cas system during infection. **a**, Tenfold dilutions of lysates of anti-CRISPR phage JBD30, and the same phage with a frameshift mutation introduced into the anti-CRISPR gene *35* (gene *35fs*), were applied to lawns of PA14 or PA14 ΔCR/*cas*. This experiment was done in a manner similar to those shown in Fig. 1a. **b**, Schematic representation of the *in vivo* homologous recombination between phage DMS3*m* and the anti-CRISPR region from phage JBD30. The 'X' approximately marks the mapped region of recombination, up- and downstream of the anti-CRISPR gene *35* from JBD30 with details shown in Supplementary Fig. 12 and Methods. **c**, Tenfold dilutions of lysates of a CRISPR-insensitive phage (DMS3), a CRISPR-sensitive phage (DMS3*m*) and DMS3*m* with anti-CRISPR gene *35* from JBD30 inserted (DMS3*m*::gene *35*) were applied to lawns of PA14 or PA14 ΔCR/*cas*. **d**, A plasmid containing a protospacer matching CR1_sp1 (shown in Fig. 1b) was electroporated into the indicated lysogens or parent strain. As indicated, the prophages within these lysogens contain either a wild-type version of anti-CRISPR gene *35*, a frameshift mutant of this gene or no anti-CRISPR gene.

The adaptive nature of the CRISPR/Cas system and the widespread occurrence of CRISPR regions in bacterial genomes suggest that this system could be the most powerful weapon possessed by bacteria to resist invasion by foreign DNA. Before our work, the only known mechanism for phages to evade CRISPR/Cas systems was by mutation[12,23], which is a low-frequency event. Here we have provided the first demonstration that the *in vivo* activity of a CRISPR/Cas system is profoundly inhibited by any one of five different 'anti-CRISPR' genes. The existence of anti-CRISPR genes is one possible explanation of how phages have continued to proliferate despite the ubiquity and potency of CRISPR/Cas systems. The possibility that anti-CRISPR genes are diverse and widespread among phages and other mobile genetic elements may account for the large diversity of CRISPR/Cas systems and the existence of multiple CRISPR/Cas system types within single bacterial strains. This proliferation of CRISPR/Cas systems may be driven by the concomitant proliferation and diversification of anti-CRISPR genes. This newly discovered arms race may have a profound effect on the evolution of both phage and bacterial genomes, and knowledge of anti-CRISPR genes will be crucial for understanding this process. The failure to detect anti-CRISPR genes until now may be due only to a lack of systematic searches using naturally functioning *in vivo* systems. Future studies to discover more anti-CRISPR genes and explain the mechanisms of their inhibition of CRISPR/Cas systems will provide new inroads for illumination of CRISPR/Cas function.

METHODS SUMMARY

Phage plaque formation assays. Overnight cultures of *P. aeruginosa* were inoculated into molten soft agar and poured onto lysogeny broth agar plates. Once solidified, 3 μl aliquots of phage lysates at tenfold dilutions were pipetted onto the agar. Plates were incubated overnight at 30 °C or 37 °C.

Lysogen construction. Phages were isolated from clinical and environmental isolates through induction with mitomycin C and subsequent plaque purification on a susceptible host. PA14 lysogens were isolated by pipetting phage lysates onto a PA14 lawn, created as described above, followed by isolation of individual colonies from the zone of clearing. Colonies were assayed for the production of the correct phage to confirm lysogeny.

Transformation efficiency assay. To construct plasmids containing protospacer sequences that are susceptible to the PA14 CRISPR/Cas system, two 42-base complementary synthetic oligonucleotides containing the protospacer were ligated to a plasmid vector, pHERD30T.

Full Methods and any associated references are available in the online version of the paper.

Received 6 August; accepted 26 October 2012.

Published online 16 December 2012.

1. Makarova, K. S. *et al.* Evolution and classification of the CRISPR-Cas systems. *Nature Rev. Microbiol.* **9,** 467–477 (2011).
2. Makarova, K. S., Grishin, N. V., Shabalina, S. A., Wolf, Y. I. & Koonin, E. V. A putative RNA-interference-based immune system in prokaryotes: computational analysis of the predicted enzymatic machinery, functional analogies with eukaryotic RNAi, and hypothetical mechanisms of action. *Biol. Direct* **1,** 7 (2006).
3. Brouns, S. J. J. *et al.* Small CRISPR RNAs guide antiviral defense in prokaryotes. *Science* **321,** 960–964 (2008).
4. Rodriguez-Valera, F. *et al.* Explaining microbial population genomics through phage predation. *Nature Rev. Microbiol.* **7,** 828–836 (2009).
5. Labrie, S. J., Samson, J. E. & Moineau, S. Bacteriophage resistance mechanisms. *Nature Rev. Microbiol.* **8,** 317–327 (2010).
6. Jore, M. M., Brouns, S. J. J. & van der Oost, J. RNA in defense: CRISPRs protect prokaryotes against mobile genetic elements. *Cold Spring Harb. Perspect. Biol.* **4,** http://dx.doi.org/10.1101/cshperspect.a003657 (2012).
7. Haurwitz, R. E., Jinek, M., Wiedenheft, B., Zhou, K. & Doudna, J. A. Sequence- and structure-specific RNA processing by a CRISPR endonuclease. *Science* **329,** 1355–1358 (2010).
8. Deltcheva, E. *et al.* CRISPR RNA maturation by trans-encoded small RNA and host factor RNase III. *Nature* **471,** 602–607 (2011).
9. Wiedenheft, B. *et al.* Structures of the RNA-guided surveillance complex from a bacterial immune system. *Nature* **477,** 486–489 (2011).
10. Westra, E. R. *et al.* CRISPR immunity relies on the consecutive binding and degradation of negatively supercoiled invader DNA by Cascade and Cas3. *Mol. Cell* **46,** 595–605 (2012).
11. Garneau, J. E. *et al.* The CRISPR/Cas bacterial immune system cleaves bacteriophage and plasmid DNA. *Nature* **468,** 67–71 (2010).
12. Barrangou, R. *et al.* CRISPR provides acquired resistance against viruses in prokaryotes. *Science* **315,** 1709–1712 (2007).
13. Yosef, I., Goren, M. G. & Qimron, U. Proteins and DNA elements essential for the CRISPR adaptation process in *Escherichia coli*. *Nucleic Acids Res.* **40,** 5569–5576 (2012).
14. Cady, K. C., Bondy-Denomy, J., Heussler, G. E., Davidson, A. R. & O'Toole, G. A. The CRISPR/Cas adaptive immune system of *Pseudomonas aeruginosa* mediates resistance to naturally occurring and engineered phages. *J. Bacteriol.* **194,** 5728–5738 (2012).
15. Marraffini, L. A. & Sontheimer, E. J. Self versus non-self discrimination during CRISPR RNA-directed immunity. *Nature* **463,** 568–571 (2010).
16. Mojica, F. J. M., Diez-Villasenor, C., Garcia-Martinez, J. & Almendros, C. Short motif sequences determine the targets of the prokaryotic CRISPR defence system. *Microbiology* **155,** 733–740 (2009).
17. Braid, M. D., Silhavy, J. L., Kitts, C. L., Cano, R. J. & Howe, M. M. Complete genomic sequence of bacteriophage B3, a Mu-like phage of *Pseudomonas aeruginosa*. *J. Bacteriol.* **186,** 6560–6574 (2004).
18. Morgan, G. J., Hatfull, G. F., Casjens, S. & Hendrix, R. W. Bacteriophage Mu genome sequence: analysis and comparison with Mu-like prophages in *Haemophilus*, *Neisseria* and *Deinococcus*. *J. Mol. Biol.* **317,** 337–359 (2002).
19. Datsenko, K. A. *et al.* Molecular memory of prior infections activates the CRISPR/Cas adaptive bacterial immunity system. *Nature Commun.* **3,** 945 (2012).
20. Cady, K. C. & O'Toole, G. A. Non-identity-mediated CRISPR-bacteriophage interaction mediated via the Csy and Cas3 proteins. *J. Bacteriol.* **193,** 3433–3445 (2011).
21. Battle, S. E., Meyer, F., Rello, J., Kung, V. L. & Hauser, A. R. Hybrid pathogenicity island PAGI-5 contributes to the highly virulent phenotype of a *Pseudomonas aeruginosa* isolate in mammals. *J. Bacteriol.* **190,** 7130–7140 (2008).
22. Cady, K. C. *et al.* Prevalence, conservation and functional analysis of *Yersinia* and *Escherichia* CRISPR regions in clinical *Pseudomonas aeruginosa* isolates. *Microbiology* **157,** 430–437 (2011).
23. Semenova, E. *et al.* Interference by clustered regularly interspaced short palindromic repeat (CRISPR) RNA is governed by a seed sequence. *Proc. Natl Acad. Sci. USA* **108,** 10098–10103 (2011).

Supplementary Information is available in the online version of the paper.

Acknowledgements We thank D. Guttman, Y.-H. Cho, K. Cady and G. O'Toole for providing *P. aeruginosa* strains and phages. We also thank K. Severinov for providing the M13 phage and *E. coli* strains required for assaying the type 1-E system. We thank J. Brumell, A. Spence and W. Navarre for reading the manuscript. We also thank D. Bona for technical assistance. This work was supported by an Operating Grant to K.L.M. (fund number MOP- 6279) and an Emerging Team Grant to A.R.D. and K.L.M. (fund number XNE86943), both of which were from the Canadian Institutes for Health Research. J.B.D. was supported by a CIHR Canada Graduate Scholarship Doctoral Award.

Author Contributions J.B.-D. designed experiments, performed experiments and wrote the manuscript, A.P. performed experiments, K.L.M. supervised experiments, and A.R.D. designed experiments and wrote the manuscript.

Author Information Phage genomes have been deposited in the National Center for Biotechnology Information under accession numbers JX434030 (JBD5), JX434031 (JBD24), JX434032 (JBD30) and JX434033 (JBD88a). Reprints and permissions information is available at www.nature.com/reprints. The authors declare competing financial interests: details are available in the online version of the paper. Readers are welcome to comment on the online version of the paper. Correspondence and requests for materials should be addressed to A.R.D. (alan.davidson@utoronto.ca).

LETTER

doi:10.1038/nature11682

Mammalian heart renewal by pre-existing cardiomyocytes

Samuel E. Senyo[1], Matthew L. Steinhauser[1], Christie L. Pizzimenti[1], Vicky K. Yang[1], Lei Cai[1], Mei Wang[4,5], Ting-Di Wu[2,3], Jean-Luc Guerquin-Kern[2,3], Claude P. Lechene[4,5] & Richard T. Lee[1,6]

Although recent studies have revealed that heart cells are generated in adult mammals, the frequency of generation and the source of new heart cells are not yet known. Some studies suggest a high rate of stem cell activity with differentiation of progenitors to cardiomyocytes[1]. Other studies suggest that new cardiomyocytes are born at a very low rate[2–4], and that they may be derived from the division of pre-existing cardiomyocytes. Here we show, by combining two different pulse-chase approaches—genetic fate-mapping with stable isotope labelling, and multi-isotope imaging mass spectrometry—that the genesis of cardiomyocytes occurs at a low rate by the division of pre-existing cardiomyocytes during normal ageing, a process that increases adjacent to areas of myocardial injury. We found that cell cycle activity during normal ageing and after injury led to polyploidy and multinucleation, but also to new diploid, mononucleate cardiomyocytes. These data reveal pre-existing cardiomyocytes as the dominant source of cardiomyocyte replacement in normal mammalian myocardial homeostasis as well as after myocardial injury.

Despite intensive research, fundamental aspects of the mammalian heart's capacity for self-renewal are still being actively debated[5,6]. Estimates of cardiomyocyte turnover range from less than 1% per year[2–4] to more than 40% per year[7]. Turnover has been reported to either decrease[3] or increase[7] with age, and the source of new cardiomyocytes has been attributed both to the division of existing myocytes[8] and to progenitors residing within the heart[9] or in exogenous niches such as bone marrow[10]. Controversy persists about the plasticity of the adult heart; in part this controversy is due to methodological challenges associated with studying slowly replenished tissues. Toxicity attributed to radiolabelled thymidine[11] and halogenated nucleotide analogues[12] limits the duration of labelling and may produce direct biological effects. Tissue autofluorescence can decrease the sensitivity and specificity of immunofluorescent methods of detecting cell cycle activity[5,13], such as probing for transient expression of cell cycle markers or the incorporation of halogenated nucleotide analogues. The challenge of measuring cardiomyocyte turnover is further compounded by the faster rate of turnover of cardiac stromal cells relative to that of cardiomyocytes[14].

Here we used multi-isotope imaging mass spectrometry (MIMS) to study cardiomyocyte turnover and to determine whether new cardiomyocytes are derived from pre-existing myocytes or from a progenitor pool (Fig. 1a). MIMS uses ion microscopy and mass spectrometry to generate high-resolution quantitative mass images and localize stable isotope reporters in domains smaller than 1 μm^3 (refs 15–17). MIMS generates ^{14}N quantitative mass images by measuring the atomic composition of the sample surface with a lateral resolution of less than 50 nm and a depth resolution of a few atomic layers. Cardiomyocyte cell borders and intracellular organelles were easily resolved (Fig. 1b). Regions of interest could be analysed at higher resolution, revealing cardiomyocyte-specific subcellular ultrastructure, including sarcomeres (Fig. 1c and Supplementary Fig. 1a). In all subsequent analyses, cardiomyocyte nuclei were identified by their location within sarcomere-containing cells, distinguishing them from adjacent stromal cells.

A great advantage of MIMS is the detection of nonradioactive stable isotope tracers. As an integral part of animate and inanimate matter, they do not alter biochemical reactions and are not harmful to the organism[18]. MIMS localizes stable isotope tracers by simultaneously quantifying multiple masses from each analysed domain; this enables the generation of a quantitative ratio image of two stable isotopes of the same element[15]. The incorporation of a tracer tagged with the rare stable isotope of nitrogen (^{15}N) is detectable with high precision by an increase in ^{15}N:^{14}N above the natural ratio (0.37%). Nuclear incorporation of [^{15}N]thymidine is evident in cells that have divided during a 1-week labelling period, as observed in the small-intestinal epithelium, which turns over completely in 3–5 days[16] (Fig. 1d); in contrast, [^{15}N]thymidine-labelled cells are rarely observed in the heart (Fig. 1e) after 1 week of labelling. In subsequent studies, small intestine was used as a positive control to confirm label delivery.

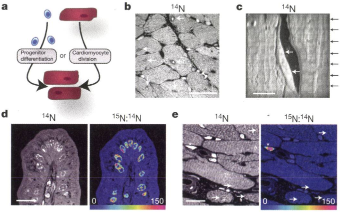


Figure 1 | Use of MIMS to study cardiomyocyte turnover. a, Primary question: are new cardiomyocytes derived from progenitors or from pre-existing cardiomyocytes? **b**, ^{14}N mass image. Subcellular details are evident, including cardiomyocyte nuclei (white arrows). Scale bar, 20 μm. **c**, MIMS resolves periodic sarcomeres (black arrows) in cardiomyocytes. Non-cardiomyocytes (white arrows) are seen outside cardiomyocyte borders. Scale bar, 5 μm. **d**, Right, ^{15}N:^{14}N hue–saturation–intensity image of small-intestinal epithelium after labelling with [^{15}N]thymidine. The scale ranges from blue, where the ratio is equivalent to natural ratio (0.37%, expressed as 0% above natural ratio (enrichment over natural ratio)), to red, where the ratio is 150% above natural ratio. ^{15}N labelling is concentrated in nuclei in a pattern resembling chromatin. Scale bar, 15 μm. **e**, Right, ^{15}N:^{14}N hue–saturation–intensity image of heart section (left ventricle). [^{15}N]Thymidine was administered for 1 week. Asterisk, rare ^{15}N$^+$ interstitial cells. Cardiomyocyte nuclei (white arrows) are unlabelled. Scale bar, 15 μm.

---

[1] Cardiovascular Division, Department of Medicine, Brigham and Women's Hospital and Harvard Medical School, Cambridge, Massachusetts 02139, USA. [2] INSERM U.759, 91405 Orsay, France. [3] Institut Curie, Laboratoire de Microscopie Ionique, 91405 Orsay, France. [4] National Resource for Imaging Mass Spectrometry, Cambridge, Massachusetts 02139, USA. [5] Genetics Division, Department of Medicine, Brigham and Women's Hospital and Harvard Medical School, Cambridge, Massachusetts 02139, USA. [6] Harvard Stem Cell Institute, Cambridge, Massachusetts 02139, USA.

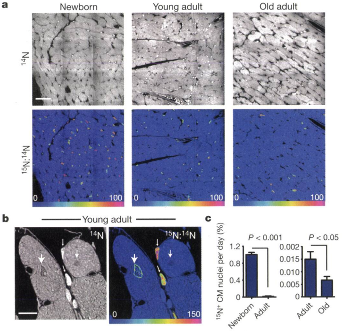


Figure 2 | Cardiomyocyte DNA synthesis decreases with age. **a**, [^{15}N]Thymidine was administered for 8 weeks to mice of different ages: newborn, starting at postnatal day 4; young adult, starting at 2 months; old adult, starting at 22 months. Top, ^{14}N mass images show histological details. Bottom, ^{15}N:^{14}N hue–saturation–intensity images show ^{15}N$^+$ nuclei. Mosaics are constructed from nine tiles, 60 µm each. Scale bar, 30 µm. **b**, High-magnification analysis shows a cardiomyocyte from the young adult with nuclear ^{15}N labelling (large arrow), two labelled non-cardiomyocytes (small arrows) and an adjacent unlabelled cardiomyocyte nucleus (medium arrow). Scale bar, 10 µm. **c**, Age-related decline in cardiomyocyte DNA synthesis. Left, comparison of newborn with young adult. Right, scale reduced to compare young adult with old adult ($n = 3$ mice per group). Error bars indicate s.e.m.

To evaluate an age-related change in cell cycle activity, we administered [^{15}N]thymidine for 8 weeks to three age groups of C57BL6 mice starting at day 4 (newborn), at 10 weeks (young adult) and at 22 months (old adult) (Supplementary Fig. 2). We then performed MIMS analysis (Fig. 2a, b and Supplementary Fig. 3). In the newborn group, 56% ± 3% (mean ± s.e.m.; $n = 3$ mice) of cardiomyocytes showed ^{15}N nuclear labelling, consistent with the well-accepted occurrence of cardiomyocyte DNA synthesis during postnatal development[19]. We observed a marked decline in the frequency of ^{15}N-labelled cardiomyocyte nuclei (^{15}N$^+$ CM) in the young adult (newborn, 1.00% ± 0.05% ^{15}N$^+$ CM per day; young adult, 0.015% ± 0.003% ^{15}N$^+$ CM per day (means ± s.e.m.); $n = 3$ mice per group, $P < 0.001$) (Fig. 2a, c and Supplementary Fig. 3). We found a further decrease in cardiomyocyte DNA synthesis in old mice (young adult, 0.015% ± 0.003% ^{15}N$^+$ CM per day; old adult, 0.007% ± 0.002% ^{15}N$^+$ CM per day (means ± s.e.m.); $n = 3$ mice per group, $P < 0.05$) (Fig. 2c). The observed pattern of ^{15}N labelling in cardiomyocyte nuclei is consistent with the known chromatin distribution pattern in cardiomyocytes[20] (Supplementary Fig. 1b) and was measured at levels that could not be explained by DNA repair (Supplementary Fig. 4). Extrapolating DNA synthesis measured in cardiomyocytes over 8 weeks yields a yearly rate of 5.5% in the young adults and 2.6% in the old mice. Given that cardiomyocytes are known to undergo DNA replication without completing the cell cycle[19,21,22], these calculations represent the upper limit of cardiomyocyte generation under normal homeostatic conditions, indicating a low rate of cardiogenesis.

To test whether cell cycle activity occurred in pre-existing cardiomyocytes or was dependent on a progenitor pool, we performed [^{15}N]thymidine labelling of double-transgenic MerCreMer/ZEG mice, previously developed for genetic lineage mapping (Fig. 3a)[23,24]. MerCreMer/ZEG cardiomyocytes irreversibly express green fluorescent protein (GFP) after treatment with 4-OH-tamoxifen, allowing pulse labelling of existing cardiomyocytes with a reproducible efficiency of about 80%. Although some have reported rare GFP expression by non-cardiomyocytes with this approach[25], we did not detect GFP expression in interstitial cells isolated from MerCreMer/ZEG hearts, nor did we detect GFP expression by Sca1 or c-Kit-expressing progenitors in histological sections (Supplementary Fig. 5). Thus, during a chase period, cardiomyocytes generated from progenitors should be GFP$^-$, whereas cardiomyocytes arising from pre-existing cardiomyocytes should express GFP at a frequency similar to the background rate induced by 4-OH-tamoxifen. We administered 4-OH-tamoxifen for 2 weeks to 8-week-old mice ($n = 4$); during a subsequent 10-week chase, mice received [^{15}N]thymidine by means of an osmotic minipump (20 µg h^{-1}). With MIMS analysis, we identified 35 ^{15}N$^+$ cardiomyocytes (of 4,190 analysed) over 10 weeks, yielding a projected yearly rate of cardiomyocyte DNA replication of 4.4%. Extrapolating from previous reports of high stem-cell-dependent cardiomyocyte turnover[7], we had expected to detect more than 320 cardiomyocytes entering the cell cycle; these results exclude such a high rate of turnover (expected ^{15}N$^+$ CM = 321; observed = 35, Fisher's exact $P < 0.0001$). Immunofluorescent staining for GFP was performed on adjacent sections, and an observer unaware of MIMS analysis results assessed GFP status. Of ^{15}N$^+$ cardiomyocytes, 77% expressed GFP, a frequency essentially identical to that of surrounding ^{15}N$^-$ cardiomyocytes (^{15}N$^+$ CM, 77%; ^{15}N$^-$ CM, 84%, Fisher's exact P = n.s.) (Table 1). If new cardiomyocytes were derived from progenitors, ^{15}N$^+$ cardiomyocytes would have been GFP$^-$ (expected = 0/35; observed = 27/35, Fisher's exact $P < 0.0001$). These data show that ^{15}N$^+$ cardiomyocytes resulted from DNA synthesis by pre-existing cardiomyocytes and exclude a substantial contribution from stem cells in cardiomyocyte replacement in the uninjured heart.

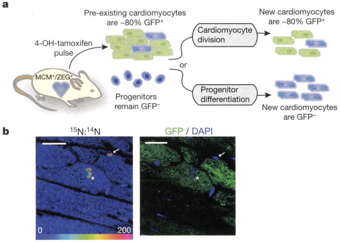

a
Pre-existing cardiomyocytes are ~80% GFP$^+$
4-OH-tamoxifen pulse
MCM$^+$/ZEG$^+$
Progenitors remain GFP$^-$
Cardiomyocyte division
or
Progenitor differentiation
New cardiomyocytes are ~80% GFP$^+$
New cardiomyocytes are GFP$^-$
b
^{15}N:^{14}N
GFP / DAPI
0
200

Figure 3 | New cardiomyocytes are derived from pre-existing cardiomyocytes during ageing. **a**, Experimental strategy. MerCreMer$^+$/ZEG$^+$ (MCM$^+$ZEG$^+$) mice ($n = 4$) were treated for 2 weeks with 4-OH-tamoxifen to induce cardiomyocyte-specific GFP expression. [^{15}N]Thymidine was administered continuously during a 10-week chase, then cycling cells were identified by ^{15}N labelling. New cardiomyocytes (^{15}N$^+$) derived from pre-existing cardiomyocytes should express GFP at a rate similar to that of the surrounding quiescent (^{15}N$^-$) cardiomyocytes. New cardiomyocytes (^{15}N$^+$) derived from progenitors should be GFP$^-$. **b**, Left, ^{15}N:^{14}N hue–saturation–intensity image showing a [^{15}N]thymidine-labelled cardiomyocyte nucleus (white asterisk) and a ^{15}N$^+$ non-cardiomyocyte (white arrow). Right, immunofluorescent image showing that the ^{15}N$^+$ cardiomyocyte is GFP$^+$. Scale bars, 15 µm.

Cardiomyocytes can undergo DNA replication without completing the cell cycle. Although multinucleation and polyploidization occur during early postnatal development and in response to myocardial

Table 1 | $^{15}N^+$ cardiomyocytes ($^{15}N^+$ CM) during normal ageing

	Total counted	GFP$^+$	Percentage GFP$^+$
$^{15}N^-$ CM	1,111	933	84
$^{15}N^+$ CM	35	27	77
Polyploid CM	12	10	83
Diploid CM	22	17	77
Multinucleate CM	18	14	78
Mononucleate CM	17	14	82
Diploid/mononucleate CM	6	5	83

The results include a new diploid/mononucleate fraction, demonstrating a similar rate of GFP expression compared with background $^{15}N^-$ cardiomyocytes ($^{15}N^-$ CM).

stress[2,19], we considered the possibility that these processes could account for $^{15}N^+$ cardiomyocytes in the uninjured adult mouse. We performed fluorescent *in situ* hybridization in adjacent sections to assess the ploidy state of each $^{15}N^+$ cardiomyocyte and surrounding $^{15}N^-$ cardiomyocytes, and an observer unaware of the results of MIMS analysis identified fluorescently labelled chromosomes (Supplementary Fig. 6). Although we found $^{15}N^+$ cardiomyocytes that were polyploid (4*n* or greater), we observed a higher frequency of diploid nuclei in the $^{15}N^+$ pool than in surrounding $^{15}N^-$ cardiomyocytes ($^{15}N^+$ diploid:polyploid = 22:12; $^{15}N^-$ diploid:polyploid = 9:56, Fisher's exact $P < 0.0001$), consistent with ongoing cell division. We then assessed each cell for multinucleation by using serial 0.5-μm sections adjacent to the section used for MIMS analysis (Supplementary Fig. 6). We observed that 49% of $^{15}N^+$ cardiomyocytes were mononucleate, in contrast with a frequency of 24% for surrounding $^{15}N^-$ cardiomyocytes (Fisher's exact $P < 0.01$), also consistent with cell division. Most cardiomyocyte DNA synthesis occurred in polyploid and/or multinucleate cardiomyocytes, as might be expected with a physiological hypertrophic response, and was thus unlikely to indicate cardiomyocyte division; however, 17% (6 of 35 $^{15}N^+$ CM) were diploid and mononucleate, consistent with newly generated cardiomyocytes (Supplementary Fig. 7). The mononucleate, diploid, $^{15}N^+$ cardiomyocytes were also predominantly GFP$^+$ (5 of 6 = 83%, compared with 82% background frequency, Fisher's exact P = n.s.), suggesting that they arose from pre-existing cardiomyocytes at a slow annual rate of 0.76% (n = 6 of 4,190 over 10 weeks).

We next used MIMS and genetic fate mapping to study myocardial injury. Cardiomyocyte GFP labelling was induced in MerCreMer/ZEG mice with 4-OH-tamoxifen. Mice then underwent experimental myocardial infarction or sham surgery followed by continuous labelling with [^{15}N]thymidine for 8 weeks. The frequency of $^{15}N^+$ cardiomyocytes in sham-operated mice was similar to that in previous experiments in unoperated mice (yearly projected rates: sham, 6.8%; unoperated, 4.4%), but increased significantly adjacent to infarcted myocardium (total $^{15}N^+$ cardiomyocyte nuclei: myocardial infarction, 23.0%; sham, 1.1%; Fig. 4a, b and Supplementary Fig. 8). We examined the nucleation, ploidy status and GFP expression of $^{15}N^+$ cardiomyocytes and surrounding unlabelled cardiomyocytes. We found a significant dilution of the GFP$^+$ cardiomyocyte pool at the border region, as previously shown[23,24] (67% versus 79%, $P < 0.05$; Table 2 and Supplementary Fig. 9); however, $^{15}N^+$ myocytes demonstrated a similar frequency of GFP expression to that in unlabelled myocytes (71% versus 67%, Fisher's exact P = n.s.), suggesting that DNA synthesis was primarily occurring in pre-existing cardiomyocytes. Of $^{15}N^+$ cardiomyocytes, about 14% were mononucleate and diploid, consistent with division of pre-existing cardiomyocytes (Supplementary Figs 6 and 7). We observed a higher DNA content (more than 2*n*) in the remaining cardiomyocytes, as expected with compensatory hypertrophy after injury. Thus, in the 8 weeks after myocardial infarction, roughly 3.2% of the cardiomyocytes adjacent to the infarct had unambiguously undergone division (total $^{15}N^+ \times$ mononucleate diploid fraction = 23% × 0.14 = 3.2%). The low rate of cardiomyocyte cell cycle completion was further supported by the absence of detectable aurora B kinase, a transiently expressed cytokinesis marker, which was detected in rapidly proliferating small intestinal cells but not in cardiomyocytes

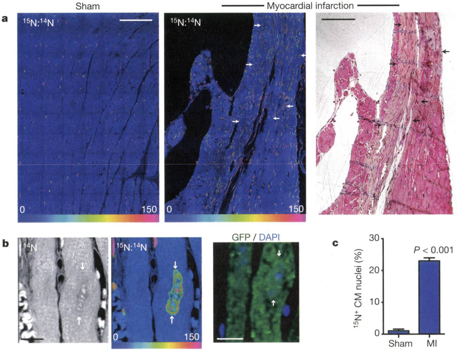


Figure 4 | Myocardial injury stimulates division of pre-existing cardiomyocytes. **a**, Myocardial infarction leads to extensive DNA synthesis within and adjacent to a scar (arrows). MerCreMer$^+$/ZEG$^+$ mice were treated for 2 weeks with 4-OH-tamoxifen to induce cardiomyocyte-specific GFP expression before myocardial infarction or sham surgery, then [^{15}N]thymidine was administered continuously for 8 weeks. Mosaics of 70 MIMS tiles (each 60 μm × 60 μm) are shown. Trichrome-stained adjacent section (right) shows the scar (black arrows). Scale bars, 90 μm. **b**, [^{15}N]Thymidine-labelled cardiomyocyte nucleus (white arrows) from myocardial infarction border region. Immunofluorescent staining shows that the cardiomyocyte is GFP$^+$. Scale bars, 10 μm. **c**, Mean percentage of $^{15}N^+$ cardiomyocyte nuclei after myocardial infarction (MI; n = 4) in the scar border region compared with that in sham-operated mice (n = 3). Error bars indicate s.e.m.

Table 2 | GFP analysis of ^{15}N^{+} cardiomyocytes (^{15}N^{+} CM)

	Total counted	GFP^{+}	Percentage GFP^{+}
Sham	3,411	2,693	79
^{15}N^{+} CM	23	19	83
Myocardial infarction	7,063	4,766	67
^{15}N^{+} CM	205	146	71
Polyploid CM	126	97	77
Diploid CM	65	47	72
Multinucleate CM	59	42	71
Mononucleate CM	41	29	71
Diploid/mononucleate CM	16	11	69

Pooled analysis of four mice in the myocardial infarction group and three mice in the sham group. Myocardial infarction led to dilution of GFP^{+} cardiomyocytes (CM) adjacent to the scar. New cardiomyocytes generated after myocardial infarction (the ^{15}N^{+}/diploid/mononucleate pool) were predominantly GFP^{+}, which is consistent with the division of pre-existing cardiomyocytes.

(Supplementary Fig. 10). We also considered the possibility that a subset of ^{15}N^{+} myocytes that were multinucleate and/or polyploid resulted from division followed by additional rounds of DNA synthesis without division. However, quantitative analysis of the ^{15}N^{+} population did not identify a subpopulation that had accumulated additional ^{15}N-label, in contrast with what would be expected in such circumstances (Supplementary Fig. 11). Taken together, these data suggest that adult cardiomyocytes retain some capacity to re-enter the cell cycle, but that most DNA synthesis after injury occurs in pre-existing cardiomyocytes without completion of cell division.

If dilution of the GFP^{+} cardiomyocyte pool cannot be attributed to division and differentiation of endogenous progenitors, do these data exclude a role for progenitors in the adult mammalian heart? The data could be explained by preferential loss of GFP^{+} cardiomyocytes after injury, a process that we have previously considered but for which we have not found supporting evidence[23]. Such an explanation excludes a role for endogenous progenitors in cardiac repair and would be consistent with data emerging from lower vertebrates[8,26] and the neonatal mouse[27] in which pre-existing cardiomyocytes are the cellular source for cardiomyocyte repletion. A second possibility to explain the dilution of the GFP^{+} cardiomyocyte pool is that injury stimulates progenitor differentiation without division; inevitably, this would lead to exhaustion of the progenitor pool, which if true could explain the limited regenerative potential of the adult mammalian heart.

This study demonstrates the birth of cardiomyocytes from pre-existing cardiomyocytes at a projected rate of roughly 0.76% per year (^{15}N^{+} annual rate × mononucleate diploid fraction = 4.4% × 0.17) in the young adult mouse under normal homeostatic conditions, a rate that declines with age but increases after myocardial injury in the border region. This study shows that cardiac progenitors do not have a significant function in myocardial homeostasis in mammals and suggests that their role after injury is also limited.

METHODS SUMMARY

Multi-isotope imaging mass spectrometry (MIMS) was performed as described previously[16]. Mice were used in accordance with the Guide for the Use and Care of Laboratory Animals and experiments were approved by the Harvard Medical School Standing Committee on Animals. Male C57Bl/6 mice were used for the ageing experiment (Charles River). Double transgenic male MerCreMer/ZEG mice[23] were used in all other experiments.

Full Methods and any associated references are available in the online version of the paper.

Received 1 December 2011; accepted 18 October 2012.

Published online 5 December 2012.

1. Hosoda, T. *et al.* Clonality of mouse and human cardiomyogenesis *in vivo*. *Proc. Natl Acad. Sci. USA* **106,** 17169–17174 (2009).
2. Soonpaa, M. H. & Field, L. J. Assessment of cardiomyocyte DNA synthesis in normal and injured adult mouse hearts. *Am. J. Physiol. Heart Circ. Physiol.* **272,** H220–H226 (1997).
3. Bergmann, O. *et al.* Evidence for cardiomyocyte renewal in humans. *Science* **324,** 98–102 (2009).
4. Walsh, S., Ponten, A., Fleischmann, B. K. & Jovinge, S. Cardiomyocyte cell cycle control and growth estimation *in vivo*—an analysis based on cardiomyocyte nuclei. *Cardiovasc. Res.* **86,** 365–373 (2010).
5. Laflamme, M. A. & Murry, C. E. Heart regeneration. *Nature* **473,** 326–335 (2011).
6. Hosoda, T., Rota, M., Kajstura, J., Leri, A. & Anversa, P. Role of stem cells in cardiovascular biology. *J. Thromb. Haemost.* **9** (Suppl. 1), 151–161 (2011).
7. Kajstura, J. *et al.* Myocyte turnover in the aging human heart. *Circ. Res.* **107,** 1374–1386 (2010).
8. Kikuchi, K. *et al.* Primary contribution to zebrafish heart regeneration by *gata4*$^{+}$ cardiomyocytes. *Nature* **464,** 601–605 (2010).
9. Beltrami, A. P. *et al.* Adult cardiac stem cells are multipotent and support myocardial regeneration. *Cell* **114,** 763–776 (2003).
10. Orlic, D. *et al.* Bone marrow cells regenerate infarcted myocardium. *Nature* **410,** 701–705 (2001).
11. Hu, V. W., Black, G. E., Torres-Duarte, A. & Abramson, F. P. ^{3}H-thymidine is a defective tool with which to measure rates of DNA synthesis. *FASEB J.* **16,** 1456–1457 (2002).
12. Wilson, A. *et al.* Hematopoietic stem cells reversibly switch from dormancy to self-renewal during homeostasis and repair. *Cell* **135,** 1118–1129 (2008).
13. Laflamme, M. A. & Murry, C. E. Regenerating the heart. *Nature Biotechnol.* **23,** 845–856 (2005).
14. Bergmann, O. *et al.* Identification of cardiomyocyte nuclei and assessment of ploidy for the analysis of cell turnover. *Exp. Cell Res.* **317,** 188–194 (2010).
15. Lechene, C. *et al.* High-resolution quantitative imaging of mammalian and bacterial cells using stable isotope mass spectrometry. *J. Biol.* **5,** 1–30 (2006).
16. Steinhauser, M. L. *et al.* Multi-isotope imaging mass spectrometry quantifies stem cell division and metabolism. *Nature* **481,** 516–519 (2012).
17. Zhang, D. S. *et al.* Multi-isotope imaging mass spectrometry reveals slow protein turnover in hair-cell stereocilia. *Nature* **481,** 520–524 (2012).
18. Klein, P. D. & Klein, E. R. Stable isotopes: origins and safety. *J. Clin. Pharmacol.* **26,** 378–382 (1986).
19. Li, F., Wang, X., Capasso, J. M. & Gerdes, A. M. Rapid transition of cardiac myocytes from hyperplasia to hypertrophy during postnatal development. *J. Mol. Cell. Cardiol.* **28,** 1737–1746 (1996).
20. Nikolova, V. *et al.* Defects in nuclear structure and function promote dilated cardiomyopathy in lamin A/C-deficient mice. *J. Clin. Invest.* **113,** 357–369 (2004).
21. Adler, C. P. & Friedburg, H. Myocardial DNA content, ploidy level and cell number in geriatric hearts: post-mortem examinations of human myocardium in old age. *J. Mol. Cell. Cardiol.* **18,** 39–53 (1986).
22. Katzberg, A. A., Farmer, B. B. & Harris, R. A. The predominance of binucleation in isolated rat heart myocytes. *Am. J. Anat.* **149,** 489–499 (1977).
23. Hsieh, P. C. *et al.* Evidence from a genetic fate-mapping study that stem cells refresh adult mammalian cardiomyocytes after injury. *Nature Med.* **13,** 970–974 (2007).
24. Loffredo, F. S., Steinhauser, M. L., Gannon, J. & Lee, R. T. Bone marrow-derived cell therapy stimulates endogenous cardiomyocyte progenitors and promotes cardiac repair. *Cell Stem Cell* **8,** 389–398 (2011).
25. Dong, F. *et al.* Myocardial CXCR4 expression is required for mesenchymal stem cell mediated repair following acute myocardial infarction. *Circulation* **126,** 314–324 (2012).
26. Jopling, C. *et al.* Zebrafish heart regeneration occurs by cardiomyocyte dedifferentiation and proliferation. *Nature* **464,** 606–609 (2010).
27. Porrello, E. R. *et al.* Transient regenerative potential of the neonatal mouse heart. *Science* **331,** 1078–1080 (2011).

Supplementary Information is available in the online version of the paper.

Acknowledgements We thank J. Gannon for surgical assistance; J. Lammerding and P. Isserman for assistance with microscopy and IP Lab software; and C. MacGillivray for histology; and J. C. Poczatek and Z. Kaufman for assistance with OpenMIMS Data Analysis and image data processing. S.E.S. is funded by the National Institutes of Health (NIH; F32 HL108570). M.L.S. is funded by the American Heart Association (AHA FTF), Future Leaders in Cardiovascular Medicine, Watkins Cardiovascular Leadership Award and the NIH (K08 DK090147). C.P.L. is funded by the NIH (EB001974 and AG034641) and the Ellison Medical Foundation (AG-SS-2215-08). R.T.L. is funded by the NIH (AG032977 and AG040019) and the Harvard Stem Cell Institute.

Author Contributions S.E.S., M.L.S. and R.T.L. designed the experiments. S.E.S., M.L.S. and V.K.Y. performed the experiments. L.C., C.L.P., V.K.Y., M.L.S. and S.E.S. performed histology. M.W. and S.E.S. operated the prototype nanoSIMS instrument. T.-D.W. operated the nanoSIMS instrument at Institut Curie with input from J.-L.G.K; S.E.S. and M.L.S. analysed the data, with input from C.P.L. and R.T.L. C.P.L. supervised all MIMS analyses and provided critical feedback at all junctures. S.E.S. and M.L.S. made the figures. S.E.S., M.L.S. and R.T.L. wrote the manuscript. All authors approved the manuscript.

Author Information Reprints and permissions information is available at www.nature.com/reprints. The authors declare no competing financial interests. Readers are welcome to comment on the online version of the paper. Correspondence and requests for materials should be addressed to R.T.L. (rlee@partners.org).

LETTER

doi:10.1038/nature11715

Structure and function of the initially transcribing RNA polymerase II–TFIIB complex

Sarah Sainsbury[1], Jürgen Niesser[1] & Patrick Cramer[1]

The general transcription factor (TF) IIB is required for RNA polymerase (Pol) II initiation and extends with its B-reader element into the Pol II active centre cleft. Low-resolution structures of the Pol II–TFIIB complex[1,2] indicated how TFIIB functions in DNA recruitment, but they lacked nucleic acids and half of the B-reader, leaving other TFIIB functions[3,4] enigmatic. Here we report crystal structures of the Pol II–TFIIB complex from the yeast *Saccharomyces cerevisiae* at 3.4 Å resolution and of an initially transcribing complex that additionally contains the DNA template and a 6-nucleotide RNA product. The structures reveal the entire B-reader and protein–nucleic acid interactions, and together with functional data lead to a more complete understanding of transcription initiation. TFIIB partially closes the polymerase cleft to position DNA and assist in its opening. The B-reader does not reach the active site but binds the DNA template strand upstream to assist in the recognition of the initiator sequence and in positioning the transcription start site. TFIIB rearranges active-site residues, induces binding of the catalytic metal ion B, and stimulates initial RNA synthesis allosterically. TFIIB then prevents the emerging DNA–RNA hybrid duplex from tilting, which would impair RNA synthesis. When the RNA grows beyond 6 nucleotides, it is separated from DNA and is directed to its exit tunnel by the B-reader loop. Once the RNA grows to 12–13 nucleotides, it clashes with TFIIB, triggering TFIIB displacement and elongation complex formation. Similar mechanisms may underlie all cellular transcription because all eukaryotic and archaeal RNA polymerases use TFIIB-like factors[5], and the bacterial initiation factor sigma has TFIIB-like topology[1,2] and contains the loop region 3.2 that resembles the B-reader loop in location, charge and function[6–8]. TFIIB and its counterparts may thus account for the two fundamental properties that distinguish RNA from DNA polymerases: primer-independent chain initiation and product separation from the template.

Our previous X-ray analysis of the Pol II–TFIIB complex at 4.3 Å resolution provided a partial backbone model of TFIIB[1]. To obtain a complete and atomic structure, we co-crystallized Pol II with a TFIIB variant lacking the mobile amino-terminal tail and carboxy-terminal cyclin fold (Methods, Supplementary Table and Supplementary Fig. 1). The resulting Pol II–TFIIB structure at 3.4 Å resolution provides details of the interactions of the four TFIIB domains with Pol II: the B-ribbon with the dock, the B-core N-terminal cyclin fold with the wall, the B-reader helix with the RNA exit tunnel, and the B-linker helix with the coiled-coil of the clamp (Fig. 1).

The structure reveals the entire course of the TFIIB polypeptide chain through the Pol II cleft, including the previously lacking[1] B-reader loop (residues 67–79) and a new 'B-reader strand' (residues 80–83). The B-reader loop does not reach the active site, but instead interacts with the Pol II rudder and fork loop 1 to seal part of the cleft. The B-reader strand adds to the two-stranded polymerase lid (Fig. 1a, c), positioning the B-reader and explaining why the lid is essential for transcription[9]. There is no evidence for a B-finger[2].

TFIIB binding induces changes in Pol II domains (Fig. 2). The lobe, protrusion and wall rotate slightly, partially closing the cleft. The previously mobile[10] tip of the protrusion forms a β-sheet and an ordered loop (α11–α12) that is required for initiation complex stability[11] (Fig. 1d). To induce these changes, the B-ribbon may bind the dock, causing the wall to rotate, which then enables the B-core to bind the wall and protrusion (Figs 1d and 2a). TFIIF also binds the lobe and protrusion[12,13], and stabilizes TFIIB on Pol II[14], in particular binding of the B-core to the wall[15]. These changes contribute to TFIIB function, because mutations in the lobe and protrusion (Supplementary Fig. 2) suppress a TFIIB mutation that causes shifts in the transcription start site (TSS)[16].

The structural changes extend to the polymerase active site, where difference electron density revealed two metal ions with similar occupancy (Fig. 2b). This was consistent with a two-metal ion catalytic mechanism[17,18], but was unexpected because free Pol II only binds metal A, whereas metal B was observed previously only when a nucleoside triphosphate (NTP) was present[19,20]. The two metal ions are bridged by a rotated aspartate D481 side chain (Fig. 2b), which is also observed in structures of NTP-containing elongation complexes of Pol II[20] and bacterial polymerase[21]. The rearrangement abolishes a contact between metal A and residue D483, and re-orientates the backbone carbonyl in the neighbouring Rpb2 residue D837 (Fig. 2b).

This indicates that TFIIB induces active-site rearrangements that increase the affinity for metal B and stimulate catalysis. Consistent with this, the B-reader of the archaeal TFIIB homologue stimulates transcription[22]; truncation of the B-reader loop in human TFIIB impairs extension of a RNA dinucleotide[23]; and the presence of TFIIB and TBP stimulates transcription initiation from pre-opened DNA[15,24]. We find here that TFIIB alone stimulates *de novo* RNA synthesis by Pol II at low or near-physiological NTP concentrations (Fig. 2c and Supplementary Fig. 3). Thus, TFIIB stimulates RNA chain initiation without contacting the catalytic site, by activating Pol II allosterically.

We also solved the structure of an initially transcribing Pol II–TFIIB complex (ITC) by soaking into the crystals a *HIS4* promoter-based scaffold with a downstream DNA duplex and a 6-base-pair (bp) DNA–RNA hybrid. Electron density was observed for the entire RNA, 7 bp of downstream DNA, and the DNA template strand until position −8 upstream of the NTP site +1. The structure at 3.6 Å resolution (Fig. 3 and Supplementary Table) shows again two metal ions bridged by a rotated D481 side chain. Metal A binds the RNA 3′ end and could bind the NTP α-phosphate[21] (Fig. 3b and Supplementary Fig. 4). Metal B is 5 Å away from metal A and may move closer for catalysis upon binding NTP β- and γ-phosphates[21] (Fig. 3b).

The ITC structure indicates that TFIIB stimulates initial RNA synthesis not only by allosteric active-site rearrangement but also by preventing tilting of short DNA–RNA hybrids. Without TFIIB, short hybrids are tilted, and this impairs synthesis, because it occludes the templating DNA base +1 from the active site[25]. Such tilting is prevented in the ITC because TFIIB binds the upstream template strand. B-reader residues R64 and D69 contact template strand bases at positions −7/−8 (Fig. 3c). These interactions may also stabilize the initial DNA bubble during open complex formation.

---

[1] Gene Center and Department of Biochemistry, Center for Integrated Protein Science CIPSM, Ludwig-Maximilians-Universität München, Feodor-Lynen-Str. 25, 81377 Munich, Germany.

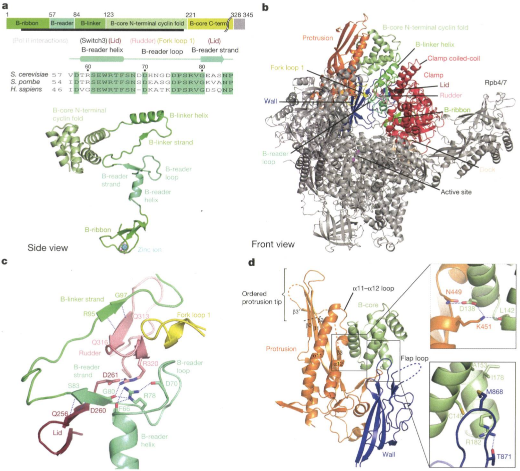

a
1
57
84
123
221
328 345
B-ribbon
B-reader
B-linker
B-core N-terminal cyclin fold
B-core C-term
(Pol II interactions)
(Switch3) (Lid)
(Rudder) (Fork loop 1)
(Lid)
B-reader helix
B-reader loop
B-reader strand
60
70
80
S. cerevisiae 57 VDTRSEWRTFSNDDHNGDDPSRVGEASNP
S. pombe 54 IDTRSEWRTFSNSDEASGDPSRVGKVANP
H. sapiens 46 IDVGSEWRTFSN-DKATKDPSRVGDSQNP
B-core N-terminal cyclin fold
B-linker helix
B-linker strand
B-reader loop
B-reader strand
B-reader helix
B-ribbon
Zinc ion
Side view
b
B-core N-terminal cyclin fold
Protrusion
B-linker helix
Clamp coiled-coil
Fork loop 1
Clamp
Lid
Rpb4/7
Wall
Rudder
B-ribbon
B-reader loop
Dock
Active site
Front view
c
B-linker strand
G97
R95
Q313
Fork loop 1
Q316
Rudder
R320
B-reader loop
B-reader strand
D261
D70
S83
G80
R78
Q256
D260
F66
Lid
B-reader helix
d
Ordered protrusion tip
β3′
β2
β0
β1
α11–α12 loop
B-core
Protrusion
β3
α11
α12
Flap loop
Wall
N449
D138
L142
K451
A153
I178
M868
C149
R182
T871

Figure 1 | Structure of the Pol II–TFIIB complex at 3.4 Å resolution. **a**, TFIIB domain organization and ribbon model. The colour code is used throughout the figures. Invariant residues are highlighted. **b**, View of Pol II–TFIIB structure showing the active centre cleft. **c**, Interactions of B-linker and B-reader with Pol II rudder, lid and fork loop 1. **d**, Interactions of B-core N-terminal cyclin fold with Pol II protrusion and wall. Previous secondary structure nomenclature was maintained[19] with newly observed elements termed β0 (Rpb2 residues 69–74), β3′ (148–151) and loop α11–α12 (438–443).

The same contacts contribute to the recognition of an initiator (Inr) DNA sequence that defines the TSS and requires a thymine at template position −7/−8 (refs 1, 26, 27). Interaction of an Inr thymine with the B-reader residue R64 (Fig. 3c) apparently occurs in the open DNA complex, because mutation of R64 causes TSS shifts that are sensitive to changes in Inr sequence[27,28]. The B-reader function in Inr recognition is supported by mutational analysis of the invariant TFIIB residue R78, which buttresses the B-reader loop, including the DNA-binding residue D69 (Fig. 1c). Mutations of R78 shift the TSS in yeast[16] and human[29], and can be lethal (Supplementary Fig. 5). Thus, the B-reader adopts a defined structure to assist in TSS selection.

The path of the RNA is blocked by the B-reader loop (Fig. 3). When the RNA grows from 6 to 7 nucleotides, it would clash with three aspartate residues (D70, D74, and D75; Fig. 3c), of which two are conserved among eukaryotes (Fig. 1a). This is consistent with a destabilization of TFIIB binding when RNA grows beyond 6 nucleotides, but TFIIB can be retained beyond this point and is only released when the RNA length reaches 12–13 nucleotides[14]. Thus, the B-reader loop

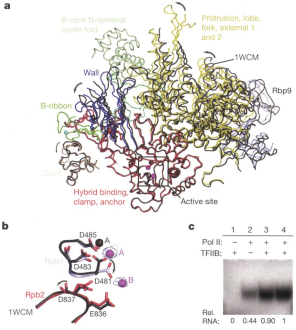

a
B-core N-terminal cyclin fold
Protrusion, lobe, fork, external 1 and 2
1WCM
Wall
Rbp9
B-ribbon
Dock
Hybrid binding, clamp, anchor
Active site
b
D485
A
A
Rpb1
D483
D481
B
Rpb2
D837
1WCM
E836
c
1 2 3 4
Pol II: − + + +
TFIIB: + − + +
Rel. RNA: 0 0.44 0.90 1

Figure 2 | TFIIB changes Pol II domains and stimulates catalysis. Same colour code as Fig. 1. **a**, Apparent order of TFIIB domain binding and polymerase domain movements. Pol II–TFIIB structure (colours) and free Pol II[10] (dark grey) were superimposed with their Rpb1 subunits. **b**, TFIIB induces active-site rearrangements. Metals A and B are shown as magenta spheres. Difference electron density was contoured at 3.0σ (grey mesh, peak height 4.4σ (metal A) and 3.6σ (metal B)). **c**, TFIIB stimulates *de novo* RNA synthesis from a *HIS4* promoter-derived tailed template. At least twofold stimulation occurs irrespective of whether TFIIB is added before (lane 3) or after (lane 4) DNA.

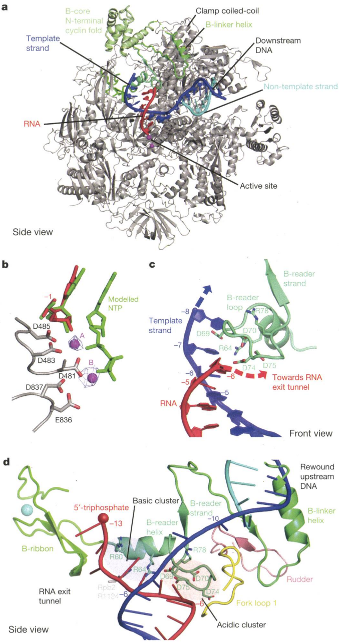


Figure 3 | Structure of initially transcribing Pol II–TFIIB complex (ITC) elucidates RNA initiation and separation, and TFIIB release. Same colour code as Fig. 1. **a**, Ribbon model. The view is related to that in Fig. 1a by a 90° rotation along a vertical axis. **b**, Active-site configuration with RNA 3′-nucleotide in position −1 (red) and superimposed 3′-nucleotide and modelled NTP (green) from a bacterial elongation complex[21]. The metals are covered by difference electron density (grey mesh) contoured at 3.0σ (peak heights 3.8σ (metal A) and 4.5σ (metal B), respectively). **c**, B-reader binds the DNA template strand at position −7/−8 and separates the RNA. **d**, Model of RNA extended into the exit tunnel elucidates TFIIB release.

separates growing RNA from the DNA template and directs it to its exit tunnel. In the elongation complex, RNA separation is achieved by rewinding of the upstream DNA duplex. However, in an ITC with a short RNA that does not yet reach the exit tunnel, such DNA rewinding probably causes collapse of the open DNA bubble, displacing RNA and aborting transcription.

Consistent with a role of the B-reader loop in RNA separation, abortive transcription increases when the loop is destabilized by R78 mutation[16] (Fig. 1c). B-reader loop function is required for transcription because deletion or mutation of B-reader loop residues 69–75 cause defects *in vivo* and *in vitro* (Supplementary Fig. 5), consistent with results for human TFIIB[23] and with *in vitro* data for yeast TFIIB[30]. B-reader loop function is conserved because replacing the loop for its *Schizosaccharomyces pombe* counterpart preserves function (Supplementary Fig. 5).

We collected diffraction data on ITCs with a 6-nucleotide RNA containing an additional 5′-phosphate and with a 7-nucleotide RNA. The data showed that the B-reader loop was unchanged, consistent with its function in RNA separation. Nucleic acids were present, but partial mobility prevented structure completion. The 6-nucleotide RNA 5′-phosphate was revealed beneath the B-reader loop, and guided modelling of an extended RNA.

The model indicates that RNA separation may be achieved by charge repulsion between the RNA 5′-triphosphate and a cluster of acidic residues in the B-reader loop, and by attraction of the triphosphate to a cluster of basic residues in the exit tunnel (Fig. 3d). Alternatively, the RNA 5′-triphosphate may be trapped by B-reader loop aspartates and a bridging metal ion. Our model accommodates 11–12 nucleotides of RNA. Further RNA extension results in a clash with the B-reader, B-ribbon and wall, explaining why TFIIB is released when the RNA length reaches 12–13 nucleotides[14]. TFIIB release enables formation of the mature 8-bp DNA–RNA hybrid observed in the elongation complex.

Bacterial RNA polymerase and its initiation factor sigma apparently use similar mechanisms for transcription initiation. The sigma factor is topologically related to TFIIB[1,2] and contains the loop region 3.2, which resembles the B-reader loop in location and negative charge[6–8]. Region 3.2 is required for formation of the first RNA phosphodiester bond, normal abortive transcription and sigma factor release[6–8].

METHODS SUMMARY

Saccharomyces cerevisiae 12-subunit Pol II (3.0 mg ml^{-1}) was incubated with a fivefold molar excess of TFIIB for 30 min at 298 K before crystallization by vapour diffusion using 32–34% (v/v) tacsimate pH 7.0, 100 mM HEPES pH 7.0, and 5 mM dithiothreitol (DTT) as reservoir solution. Crystals were cryo-protected in mother solution containing 70% (v/v) tacsimate, supplemented with 670 μM *HIS4* tailed template[25] for ITC complexes, and incubated at 281 K over night before flash-cooling in liquid nitrogen. Diffraction data were collected at 100 K at SLS beamline X06SA. Structures were solved by molecular replacement (Protein Data Bank code 1WCM). TFIIB was expressed from a pET21 plasmid in *Escherichia coli* strain Rosetta DE3 (Novagen) and purified by nickel affinity, cation exchange and size-exclusion chromatography. For *in vivo* analysis, TFIIB mutants were cloned into vector pRS315 containing 500 bp upstream and 300 bp downstream of the *SUA7* open reading frame (ORF), and selected on 5-FOA medium[1]. For *in vitro* transcription, nuclear extracts were prepared from wild-type and strain SHY245 (ref. 4) as described and assays performed as reported[1]. For the *de novo* transcription assay, 20 μM rCTP, rGTP and rUTP, and 2 μM rATP supplemented with 5 μCi [α-^{32}P] rATP were incubated with 2.5 pmol Pol II, 25 pmol TFIIB and 5 pmol template for 5 min at 301 K.

Full Methods and any associated references are available in the online version of the paper.

Received 27 July; accepted 26 October 2012.

Published online 14 November 2012.

1. Kostrewa, D. *et al.* RNA polymerase II–TFIIB structure and mechanism of transcription initiation. *Nature* **462,** 323–330 (2009).
2. Liu, X., Bushnell, D. A., Wang, D., Calero, G. & Kornberg, R. D. Structure of an RNA polymerase II-TFIIB complex and the transcription initiation mechanism. *Science* **327,** 206–209 (2010).
3. Cho, E. J. & Buratowski, S. Evidence that transcription factor IIB is required for a post-assembly step in transcription initiation. *J. Biol. Chem.* **274,** 25807–25813 (1999).
4. Ranish, J. A., Yudkovsky, N. & Hahn, S. Intermediates in formation and activity of the RNA polymerase II preinitiation complex: holoenzyme recruitment and a postrecruitment role for the TATA box and TFIIB. *Genes Dev.* **13,** 49–63 (1999).
5. Vannini, A. & Cramer, P. Conservation between the RNA polymerase I, II, and III transcription initiation machineries. *Mol. Cell* **45,** 439–446 (2012).
6. Vassylyev, D. G. *et al.* Crystal structure of a bacterial RNA polymerase holoenzyme at 2.6 Å resolution. *Nature* **417,** 712–719 (2002).
7. Murakami, K. S., Masuda, S. & Darst, S. A. Structural basis of transcription initiation: RNA polymerase holoenzyme at 4 Å resolution. *Science* **296,** 1280–1284 (2002).

8. Kulbachinskiy, A. & Mustaev, A. Region 3.2 of the sigma subunit contributes to the binding of the 3′-initiating nucleotide in the RNA polymerase active center and facilitates promoter clearance during initiation. *J. Biol. Chem.* **281,** 18273–18276 (2006).
9. Naji, S., Bertero, M. G., Spitalny, P., Cramer, P. & Thomm, M. Structure-function analysis of the RNA polymerase cleft loops elucidates initial transcription, DNA unwinding and RNA displacement. *Nucleic Acids Res.* **36,** 676–687 (2008).
10. Armache, K. J., Mitterweger, S., Meinhart, A. & Cramer, P. Structures of complete RNA polymerase II and its subcomplex, Rpb4/7. *J. Biol. Chem.* **280,** 7131–7134 (2005).
11. Grünberg, S., Warfield, L. & Hahn, S. Architecture of the RNA polymerase II preinitiation complex and mechanism of ATP-dependent promoter opening. *Nature Struct. Mol. Biol.* **19,** 788–796 (2012).
12. Chen, Z. A. *et al.* Architecture of the RNA polymerase II-TFIIF complex revealed by cross-linking and mass spectrometry. *EMBO J.* **29,** 717–726 (2010).
13. Eichner, J., Chen, H. T., Warfield, L. & Hahn, S. Position of the general transcription factor TFIIF within the RNA polymerase II transcription preinitiation complex. *EMBO J.* **29,** 706–716 (2010).
14. Cabart, P., Ujvari, A., Pal, M. & Luse, D. S. Transcription factor TFIIF is not required for initiation by RNA polymerase II, but it is essential to stabilize transcription factor TFIIB in early elongation complexes. *Proc. Natl Acad. Sci. USA* **108,** 15786–15791 (2011).
15. Fishburn, J. & Hahn, S. Architecture of the yeast RNA polymerase II open complex and regulation of activity by TFIIF. *Mol. Cell. Biol.* **32,** 12–25 (2012).
16. Chen, B. S. & Hampsey, M. Functional interaction between TFIIB and the Rpb2 subunit of RNA polymerase II: implications for the mechanism of transcription initiation. *Mol. Cell. Biol.* **24,** 3983–3991 (2004).
17. Steitz, T. A., Smerdon, S. J., Jager, J. & Joyce, C. M. A unified polymerase mechanism for nonhomologous DNA and RNA polymerases. *Science* **266,** 2022–2025 (1994).
18. Sosunov, V. *et al.* Unified two-metal mechanism of RNA synthesis and degradation by RNA polymerase. *EMBO J.* **22,** 2234–2244 (2003).
19. Cramer, P., Bushnell, D. A. & Kornberg, R. D. Structural basis of transcription: RNA polymerase II at 2.8 angstrom resolution. *Science* **292,** 1863–1876 (2001).
20. Wang, D., Bushnell, D. A., Westover, K. D., Kaplan, C. D. & Kornberg, R. D. Structural basis of transcription: role of the trigger loop in substrate specificity and catalysis. *Cell* **127,** 941–954 (2006).
21. Vassylyev, D. G. *et al.* Structural basis for substrate loading in bacterial RNA polymerase. *Nature* **448,** 163–168 (2007).
22. Werner, F. & Weinzierl, R. O. Direct modulation of RNA polymerase core functions by basal transcription factors. *Mol. Cell. Biol.* **25,** 8344–8355 (2005).
23. Tran, K. & Gralla, J. D. Control of the timing of promoter escape and RNA catalysis by the transcription factor IIb fingertip. *J. Biol. Chem.* **283,** 15665–15671 (2008).
24. Pal, M., Ponticelli, A. S. & Luse, D. S. The role of the transcription bubble and TFIIB in promoter clearance by RNA polymerase II. *Mol. Cell* **19,** 101–110 (2005).
25. Cheung, A. C., Sainsbury, S. & Cramer, P. Structural basis of initial RNA polymerase II transcription. *EMBO J.* **30,** 4755–4763 (2011).
26. Zhang, Z. & Dietrich, F. S. Mapping of transcription start sites in *Saccharomyces cerevisiae* using 5′ SAGE. *Nucleic Acids Res.* **33,** 2838–2851 (2005).
27. Kuehner, J. N. & Brow, D. A. Quantitative analysis of *in vivo* initiator selection by yeast RNA polymerase II supports a scanning model. *J. Biol. Chem.* **281,** 14119–14128 (2006).
28. Bangur, C. S., Pardee, T. S. & Ponticelli, A. S. Mutational analysis of the D1/E1 core helices and the conserved N-terminal region of yeast transcription factor IIB (TFIIB): identification of an N-terminal mutant that stabilizes TATA-binding protein-TFIIB-DNA complexes. *Mol. Cell. Biol.* **17,** 6784–6793 (1997).
29. Hawkes, N. A. & Roberts, S. G. The role of human TFIIB in transcription start site selection *in vitro* and *in vivo*. *J. Biol. Chem.* **274,** 14337–14343 (1999).
30. Yang, C. & Ponticelli, A. S. Evidence that RNA polymerase II and not TFIIB is responsible for the difference in transcription initiation patterns between *Saccharomyces cerevisiae* and *Schizosaccharomyces pombe*. *Nucleic Acids Res.* **40,** 6495–6507 (2012).

Supplementary Information is available in the online version of the paper.

Acknowledgements We thank C. Bäjen, A. Cheung, S. Etzold, F. Hög, D. Kostrewa, N. Pirkl and other members of the Cramer laboratory. Part of this work was performed at the Swiss Light Source (SLS) at the Paul Scherrer Institut, Villigen, Switzerland. S.S. was supported by a postdoctoral fellowship from the Alexander-von-Humboldt Foundation. P.C. was supported by the Deutsche Forschungsgemeinschaft, SFB646, TR5, GraKo1721, SFB960, CIPSM, NIM, an Advanced Grant of the European Research Council, the LMUinnovativ project Bioimaging Network, the Jung-Stiftung, and the Vallee Foundation.

Author Contributions S.S. carried out experiments except for functional analysis presented in Supplementary Fig. 5, which was carried out by J.N. P.C. initiated and supervised the project. S.S. and P.C. prepared the manuscript.

Author Information Coordinates and structure factors of the Pol II–TFIIB complex and the ITC with 6-nucleotide RNA have been deposited in the Protein Data Bank under accession numbers 4BBR and 4BBS, respectively. Reprints and permissions information is available at www.nature.com/reprints. The authors declare no competing financial interests. Readers are welcome to comment on the online version of the paper. Correspondence and requests for materials should be addressed to P.C. (cramer@lmb.uni-muenchen.de).

CAREERS

Q&A NIH director lays out workforce initiatives for early-career scientists **p.443**

NATUREJOBS BLOG Grab the latest science-careers news and issues **go.nature.com/z8g4a7**

NATUREJOBS For the latest career listings and advice **www.naturejobs.com**

BJORN RUNE LIE/GETTY

EDUCATION

Free-range learning

The budding field of informal science education offers varied research paths but uncertain funding.

BY VIRGINIA GEWIN

Zombies are on the attack. Victims' organs must be replenished to save humanity from the undead. But only players with an understanding of cells, tissues and organs — even the latest techniques to induce pluripotent stem cells — can regenerate the lost body parts.

This is the challenge set forth in *Progenitor X*, a problem-solving video game in which players learn how to reprogram more than 200 human cell types. "People play *Progenitor X* because it is fun to learn some things about science as they play," says game developer Kurt Squire, who works at the University of Wisconsin–Madison.

Squire is one of a growing number of informal-science-education (ISE) researchers, who focus on providing learning experiences outside of the formal classroom environment. The field encompasses a diverse group of talents, including video-game developers, as well as people who design, develop and evaluate science-education experiences for museums, television shows, after-school programmes, science festivals and national parks. "The learning that takes place in museums and science centres is the same learning as in school — brains generally work the same no matter the location," says Justin Dillon, an ISE researcher at Kings College London, which is launching a bachelor's degree in science engagement and communication in September. "But those experiences can reinforce or deepen the learning."

ISE is increasingly gaining traction: both government and non-profit funders are supporting activities; some museums are hiring more research staff; and the number of training opportunities is increasing. Attempting to organize the field is the Washington DC-based Center for Advancement of Informal Science Education (CAISE), which was formed in 2007 with support from the US National Science Foundation (NSF). CAISE aims to raise the field's profile by bringing together diverse sectors under the ISE umbrella to garner more funding, formalize training opportunities and establish methods to assess the success of various outreach projects.

However, in the face of strained science budgets, this budding field has been prone to funding ups and downs. Government agencies might perceive their activities as lower priorities. Practitioners have increasingly forged partnerships with industry or private foundations that are eager to cultivate a science-savvy citizenry and workforce. Although diverse job opportunities exist, seizing them requires connections, flexibility and creativity.

ACADEMIC RESEARCH

The diversity of career tracks may be the field's biggest perk. Science centres, zoos, aquaria, video-game companies, television shows and non-profit organizations are all looking to develop new materials to keep visitors and users engaged. Research-oriented jobs are among the fastest-growing areas of employment. "Funders want people to fundamentally look at the learning question — it's not just about building an exhibit any more to get NSF funding," says Rhiannon Crain, who trained with Dillon and is now leading a citizen-science project called Yard Map. Run from the Cornell Lab of Ornithology in Ithaca, New York, Yard Map asks users to provide data about the effect of landscape changes on wildlife, and teaches them how best to create a bird habitat.

ISE practitioners might quantify, for example, how access to scientific instruments or engaging in hands-on exploration — through laboratory experiments or forensic-investigation simulations — can spark curiosity and strengthen someone's comprehension of science. Kevin Crowley, an education researcher at the University of Pittsburgh in Pennsylvania, has partnered with the Carnegie Museum of Natural History in Pittsburgh to help the ▶

museum to enrich civic debates that involve scientific evidence. For example, to increase climate-change literacy in urban communities, the museum is breaking down the huge, complex topic into small pieces, focusing on how climate change will affect Pittsburgh itself.

Approaches to ISE, and therefore to training options, differ. A group led by Lynn Dierking at Oregon State University in Corvallis focuses on what motivates people to seek learning experiences in daily life, a concept she calls "free-choice learning". Dierking's group studies how to create opportunities that pique human curiosity, such as during visits to state parks.

The university offers a six-course certificate programme, an online master's degree and a formal PhD. Trainees can have a science or social-science background. The online master's is designed to provide the skills to enhance visitor learning at museums, national parks or nature centres. PhD candidates typically conduct research at an ISE outlet (such as a museum or an after-school programme), and recipients generally go into academia or become evaluators.

But the academic track often offers little job security. "Unfortunately, most of those are not tenure-track positions," says Crain. Those who do find university-based work are typically adjunct faculty who rely heavily year-to-year on grant money. Squire's students end up not only in academia but also in the entertainment industry or starting their own companies.

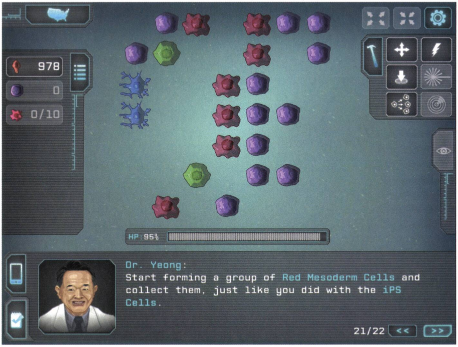
OUTSIDE ACADEMIA

Museums, meanwhile, particularly larger ones, are building up internal research and evaluation teams to track, for example, visitor experiences through surveys. The Science Museum of Minnesota in St Paul and the Exploratorium in San Francisco, California, have teams of 12 to 14 who use this visitor feedback to hone exhibits. Larger museums have programme or exhibit departments — many of which are staffed by scientists with doctoral degrees — that develop and design exhibits, a process that can take two to three years from conception to completion. Smaller museums often have to hire external evaluators or companies to design and fabricate their exhibits. (Although some museums, notably The Field Museum in Chicago, Illinois, are facing harsh budget realities that have meant cuts affecting researchers and other employees (see *Nature* http://doi.org/j6q; 2012).)

> ***"We are really interested in understanding what motivates and sustains interest in science."***

The increasing demand for consultant evaluators is a bright spot that seems poised to continue. "The need for evaluators in ISE is growing even when, and perhaps because of, decreases in funding," says Kirsten Ellenbogen, senior director of lifelong learning at the Science Museum of Minnesota. It's more urgent to demonstrate impacts than in past years, she adds. Evaluators collect data about the learning experiences of museum visitors to show whether the programme or exhibit has effectively educated or enlightened. Evaluation is often crucial to satisfying grant requirements. "In the current funding environment, accountability is critical," agrees Ellen McCallie, programme director with NSF's division of research on learning in Arlington, Virginia.

As a result, evaluators can find work as project-based consultants or join evaluation firms that are popping up. The research institute SRI International, headquartered in Menlo Park, California, continues to grow — hiring people with social-science skills and a passion for communicating science. A few universities, notably Boston College in Massachusetts, the University of California, Berkeley, and Claremont Graduate University in California, offer PhD-level training that specializes in designing quantitative-research studies to analyse visitor data.



Dr. Yeong:
Start forming a group of Red Mesoderm Cells and collect them, just like you did with the iPS Cells.

The zombie video game *Progenitor X* helps players to learn how cell types are formed.

WISCONSIN INST. DISCOVERY

FUNDING FLUCTUATIONS

Regardless of the track, funding is a challenge. The NSF cut its funding request for the Advancing Informal STEM (Science, Technology, Engineering and Maths) Learning programme for the fiscal year 2013 by 22% from 2012 numbers, to US$43 million. Funding is equally uncertain outside the United States. It is not yet clear whether 'Science in Society' — one of the two main ISE funding schemes under the European Framework Programme — will continue to exist in 2014.

Yet, at the same time, funders such as the NSF broadly encourage the use of ISE outreach and evaluation techniques across other NSF programmes and centres. For example, the NSF's Centers for Chemical Innovation programme requires a robust effort in informal science communication or ISE. The programme's director, Katharine Covert, says that there is a supplemental funding opportunity of up to $150,000 for the centres to develop collaborations with an ISE organization. And a new NSF-funded initiative — Science Education For New Civic Engagements and Responsibilities (SENCER) — will establish partnerships between higher education and ISE organizations to find innovative ways to share research and interact with the public. The next round of SENCER funding, as part of a "Civic Engagement Partnership" programme, will provide six $50,000 partnership awards (applications are due by 15 March) to ISE–higher-education collaborations.

"These days, informal-science educators have to be flexible and imaginative about how to obtain funding in this research area; ideally, they are able to piggyback onto existing activities," says Karen Bultitude, director of research in the department of science and technology studies at University College London. She expects that the UK's new Research Excellence Framework, which requires researchers to demonstrate the impact of their research outside academia, will encourage other groups to collaborate with ISE outreach projects. "Almost all of academia in the UK is desperately interested in demonstrating impacts of their work on wider groups," says Bultitude.

In some cases, private funders have stepped in. The Wellcome Trust in London has spent upwards of £50 million (US$80 million) over the past decade on schemes that include media fellowships for scientists and, most recently, engagement fellowships to support senior researchers who are developing a programme or projects to share their expertise with the public. One recipient, Roger Kneebone, based at Imperial College London, is a surgeon-turned-public-engagement fellow who plans to use live surgical simulations to engage the public. Clare Matterson, director of medical humanities and engagement at the Wellcome Trust, hopes that two reports published by the trust in November 2012 — *Analysing the UK Science Education Community: the Contribution of Informal Providers* and *Review of Informal Science Learning* — will prompt more grant submissions in the United Kingdom, and for ISE in general, of which there is a dearth at the moment, she says. “What is needed is a greater number of high-quality research-based proposals so we have a better understanding of how children learn through informal science experiences,” she adds. “These don’t have to be education groups; they could be people working in psychology or neuroscience, studying how people respond to different types of learning.”

In the United States, the Gordon and Betty Moore Foundation and the Noyce Foundation, both in Palo Alto, California, are funding ISE efforts ranging from citizen-science engagement to helping teachers in training learn how to develop their own interactive science-lesson plans. “We are really interested in understanding what motivates and sustains interest in science,” says Janet Coffey, a programme officer at the Gordon and Betty Moore Foundation.

Industry is also experimenting with ISE in an effort to reach the public. Novo Nordisk, headquartered in Bagsvaerd, Denmark, for example, finances the Steno Health Promotion Centre in Gentofte, Denmark, which runs a 33-million-krone (US$5.7 million) project called PULSE at the Experimentarium in Copenhagen, to promote healthy lifestyles for families. The project will engage and track families with children aged 6–12 years old from a variety of socioeconomic backgrounds as they design and develop healthful eating and exercise strategies. The museum is increasingly “looked at as a serious partner instead of simply an institute for kids,” says Sheena Laursen, director of international projects at the Experimentarium.

Many hope that continued research efforts that document how ISE increases scientific literacy or encourages students to pursue STEM careers will help the field to grow. “Anyone who cares about scientific literacy and STEM workforce development,” says Ellenbogen, “should pay attention to ISE.” ■

Virginia Gewin *is a freelance writer in Portland, Oregon.*

NIH

Q&A Francis Collins

Changes at the NIH

For years, the US National Institutes of Health has struggled with promoting non-academic career tracks for biomedical scientists, gauging the supply of PhD holders and demand for research jobs, enticing under-represented minorities into science and establishing funding avenues for early-career researchers. Hoping to bring some evidence-based clarity to these issues, NIH director Francis Collins asked two working groups of the NIH Advisory Committee to study the issues and make recommendations. They released their recommendations in two reports in June; Collins responded in December. The NIH has decided to take measures that include raising its postdoc stipend, increasing the number of grants that encourage early-career independence and offering 25 institutional grants, each worth about US$250,000, to support training programmes that prepare students for a broad range of research-related careers, including non-academic paths.

What did you learn from the advisory reports?
The working groups put in a lot of effort collecting data to figure out what career paths PhDs were ultimately heading down. Only about 23% of US-trained biomedical PhD holders were in academic tenure or tenure-track positions in 2008, they found. Many end up in research-related positions, in industry, government, teaching, science policy, science journalism and other science-related professions. Training programmes have tended to view those tracks as secondary and have even sent messages that they are second-rate. So we need to retool to expose trainees to multiple pathways, rather than simply producing clones of their principal investigators.

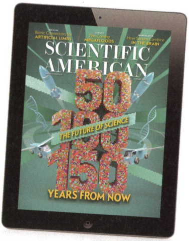
What might the $250,000 institutional grants involve?
We propose to have institutions provide actual exposure to careers outside a university setting, to help them to think more creatively about the skills and knowledge that PhD holders might need for the job market. For example, students and postdocs may spend time at a pharmaceutical company, develop entrepreneurial skills or learn aspects of science-policy development.

Where will you get the money to support these programmes?
We can do it with the NIH Common Fund, which is dedicated to innovative initiatives, especially for topics that are relevant to multiple NIH institutes or centres. We will have to balance that with the need to continue other programmes through the Common Fund.

Unemployment is low among scientists in the United States, at 4–5%. Is the priority, then, to change expectations among those considering the academic track?
Yes, that is the intention. Unemployment is actually about 2% among biomedical PhD ▶

PROMOTION

Tap into the future.



Get closer to the leading edge of science and technology: experience the SCIENTIFIC AMERICAN Tablet Edition - available for iPad®.

Apple and iPad are trademarks of Apple Inc., registered in the U.S. and other countries. App Store is a service mar

Available on the App Store

TO RENEW
SCIENTIFICAMERICAN.COM/UPGRADE

TO SUBSCRIBE
SCIENTIFICAMERICAN.COM/TABLET

holders. But although the training is aimed almost entirely at preparing people for tenure-track positions in academia, only a minority of US-trained biomedical researchers end up there.

The working groups recommend supporting students and postdocs through trainee grants to shorten the time it takes to get a degree and improve the experience. Why have you declined to do so?

It is all about logistics. Something like 80% of postdocs and two-thirds of graduate students and trainees are associated with research grants — how do we put them on training grants instead? The shift would be an enormous administrative challenge. Instead, we are asking institutions to apply training-focused principles to all students and postdocs. For example, we recommend that all trainees have individual development plans to track how they are doing.

Do you think that it is the NIH's job to decide how many biomedical PhDs are needed?

This question fascinates and troubles me. It seems as though not a week goes by without someone saying that there are too many PhDs and someone else saying there are not enough. They can't both be right.

We are planning to establish a sophisticated economic-modelling unit at the NIH to determine supply and demand. We are missing a lot of data. We need to determine who is interested in going into PhD and medical doctorate–PhD programmes, and estimate the current and future demand for their skills in all sectors.

Does science need another established career track? What would it look like?

The working groups looked at career tracks for people who don't see themselves as the head of a lab. For example, on the NIH intramural campus in Bethesda, Maryland, we have a number of staff scientists. I have two in my own laboratory. I think the output per dollar spent for these scientists is very high. The working groups recommended that the peer-review process should recognize the contributions of staff scientists and consider them more favourably. At the NIH, we can encourage study sections to recognize the value of these scientists.

What is the 'diversity challenge' that the biomedical community is facing?

The diversity of the US biomedical research workforce is woefully short of that in the overall population. That is damaging our ability to carry out research as effectively as we would like. For example, there is evidence that individuals from under-represented groups tend to be interested in disparities in health between demographic groups, because they often come from communities that have experienced them. The absence of effective programmes to recruit and retain such individuals leaves a gap in the biological sciences.

The NIH and other agencies have had various diversity initiatives, with limited effect. Can the NIH effect a sea change?

A sea change has been proposed by the diversity working group. On entrance to college, under-represented groups show a lot of interest in science. But we are losing a lot of talented people during the undergraduate years. Much of that is due to the lack of exposure to real research experience. Financial concerns about a long training period are also a factor.

We have proposed a programme called BUILD (Building Infrastructure Leading to Diversity), coming in the next 18 months. Undergraduate students at institutions that are not traditionally research-intensive will have an opportunity to do summer research in their chosen field and have their junior and senior years at college paid for. They will pay this back with two years of research in a full-time job.

The NIH is proposing to launch a mentoring network. What is its aim?

The idea is to give young scientists around the country a chance to be mentored by people who have similar interests and backgrounds. We will make awards to organizations to run the network after a competitive process.

According to a study published in 2011, African-Americans have lower rates of success with grant applications than other applicants. What is being done to address this?

We are planning more peer-review experiments, following up on concerns about possible unconscious bias in science [based on the findings in D. K. Gunther *et al. Science* 333, 1015–1019 (2011)]. We at the NIH found the evidence alarming and we are determined to get to the bottom of it.

Do you discuss career paths with the postdocs in your lab?

We talk about career paths regularly, but I'm not sure that we cover all of the options effectively. This is perhaps a wake-up call for me. Most of my trainees end up in academia, but I probably haven't done a great job of exposing them to other options.

How would you advise a postdoc struggling with career options?

It is hard to generalize; you need to take into account their dreams, their talents and the contributions that they can make in different settings. I worry that a number of them are receiving the message that if they don't get a tenure-track position, they have failed. The good news is that nearly all postdocs are likely to be employed in interesting positions, but many will not travel a narrow academic path. ■

INTERVIEW BY GENE RUSSO

naturejobs

European Head Office, London	naturejobs@nature.com
The Macmillan Building, 4 Crinan Street, London NI 9XW, UK Tel: +44 (0) 20 7843 496I Fax: +44 (0) 20 7843 4996	
Global Head: Nils Moeller,+44 (0) 20 7843 4953	n.moeller@nature.com
Southwest UK/RoW: Emma Hare, +44 (0) 20 7843 4944	e.hare@nature.com
Northeast UK/Ireland: Matthew Ward, +44 (0) 20 7014 4059	m.ward@nature.com
France/Switzerland/Belgium: Julie Aramburu, +44 (0) 20 7843 4994	j.aramburu@nature.com
Scandinavia/Spain/Portugal/Italy: Megan Ball, +44 (0) 20 7843 4973	m.ball@nature.com
North Germany/The Netherlands/Eastern Europe: Kristin Doelling, +44 (0) 20 7843 4970	k.doelling@nature.com
South Germany/Austria: Maria Kubalova, +44 (0) 20 7014 4084	m.kubalova@nature.com
Natureevents: Matthew Clare, +44 (0) 20 7014 4015	m.clare@nature.com
Business Development Manager, Middle East: Jon Giuliani, +44 (0) 20 7843 4922	j.giuliani@nature.com
USA Head Office, New York	naturejobs@us.nature.com
75 Varick Street, 9th Floor, New York, NY 10013-1917 Tel: +1 800 989 7718 Fax: +1 800 989 7103	
US Sales Manager: Ken Finnegan, +1 212 726 9248	k.finnegan@us.nature.com
Japan Head Office, Tokyo	
Chiyoda Building, 2-37 Ichigayatamachi, Shinjuku-ku, Tokyo 162-0843 Tel: +81 3 3267 8751, Fax: +81 3 3267 8746	
Asia-Pacific Sales & Business Development Manager: Yuki Fujiwara, +81-3-3267-8765	y.fujiwara@natureasia.com
Business Development Manager, Greater China/Singapore: Gloria To, +852 2811 7191	g.to@natureasia.com
India	
3a, 4th Floor, DLF Corporate Park, Gurgaon, 122 002, India	
Advertisement Sales: Vikas Chawla, +91 1242881057	v.chawla@nature.com

---

NIDCD

Department of Health and Human Services
National Institutes of Health
National Institute on Deafness and Other Communication Disorders
Tenure-Track Investigator Position

The Division of Intramural Research, National Institute on Deafness and Other Communication Disorders (NIDCD), located in Bethesda, MD, is seeking a tenure-track scientist to establish an independent research program in hearing and balance. We welcome applications from candidates with clinical or basic research interests in hearing and balance or disorders of these processes. Preference will be given to candidates whose experimental approaches complement those of our existing strong programs in the genetics, development and cell biology of hearing and balance. The successful candidate will join a dynamic group of scientists in a growing intramural program at the forefront of research on communication disorders. Research space is located in a newly constructed, inter-Institute neuroscience research center (Porter Neuroscience Research Center) which includes state-of-the art NIDCD core facilities. The NIDCD offers an exceptional working environment including well-equipped research laboratories, the NIH Clinical Research Center, and numerous opportunities for collaboration. This position includes a generous start-up allowance, an ongoing commitment of research space, laboratory resources, and positions for personnel and trainees. Candidates must possess a Ph.D. and/or M.D., post-doctoral experience, and an outstanding publication record. Salary is commensurate with education and experience.

Please submit a curriculum vitae including bibliography, three reprints of recent relevant publications, statement of research interests, an outline of your proposed research, and the names and addresses of three references to: Ms. Linda de Iberri, Office of the Scientific Director, NIDCD, 5 Research Court, Room 2B28, Rockville, MD 20892 (deiberril@nidcd.nih.gov). Applications will be reviewed starting March 1, 2013 and accepted until the position is filled. DHHS and NIH are Equal Opportunity Employers and encourage applications from women and minorities.

NW232679R

MANCHESTER
1824
The University of Manchester

Faculty of Life Sciences

The Healing Foundation Centre

Chair in Skin Repair and Regeneration

Salary will be commensurate with experience on the Professorial Salary Scale **Ref: LSX-02209**

The University of Manchester seeks to appoint an internationally renowned clinical translational scientist as Chair in Skin Repair and Regeneration within The Healing Foundation Centre for Tissue Regeneration. The post will be held within the Faculty of Life Sciences (FLS), with strong translational links with the Faculty of Medical and Human Sciences (FMHS). This appointment is an unrivalled opportunity to build on the outstanding potential in Manchester to translate basic discoveries towards skin regeneration, and to make a real difference to regenerative medicine and to delivering improved healthcare.

The successful individual, who is likely to be clinically and scientifically qualified, will lead the clinical translational activities of The Healing Foundation Centre for Tissue Regeneration within FLS. You will direct your own world class research programme in the area of wound healing and skin repair, develop cross-Faculty collaborations with scientists and clinicians, build on Manchester's internationally leading strengths in developmental, cell-extracellular matrix, stem/progenitor cell and inflammation biology, and provide leadership in wound repair and its translation into clinical medicine.

Informal enquiries can be made to Professor Cay Kielty.
Email Cay.Kielty@manchester.ac.uk

For further information and to apply, please visit our website.

If you are unable to apply online, please request an application form by emailing hrrecruitment@manchester.ac.uk

Closing date: 20 February 2013.
Please quote reference number.

The University of Manchester values a diverse workforce and welcomes applications from all sections of the community

www.manchester.ac.uk/jobs

POSITIVE ABOUT DISABLED PEOPLE

U232769R

THE CHINESE UNIVERSITY OF HONG KONG

Applications are invited for:-

School of Biomedical Sciences, Faculty of Medicine
Professor

(Ref. 1213/112(540)/2) (Closing date: March 31, 2013)

The School of Biomedical Sciences was established under the Faculty of Medicine in June 2009 through re-organization of its pre-clinical Departments. It is the Faculty's vision to become a world-recognized leader in selected fields of biomedical sciences while maintaining excellence in postgraduate and undergraduate educational programmes. Currently, the School has research programmes focusing on neuro-degeneration, neuro-development and repair, cancer biology and inflammation, developmental biology, reproduction and endocrinology, vascular and metabolic biology, and stem cell and regeneration.

Individuals who are strongly motivated by the biological importance of their research and who value the opportunity to work in close collaboration with others are welcome to apply. The School will provide state-of-the-art research facilities, appropriate laboratory and office space, and a handsome start-up package commensurate with the appointees' experience and track record.

Applicants should have (i) a PhD, MD, or DVM degree or equivalent in biological and biomedical sciences or related disciplines; (ii) expertise in using cutting-edge approaches for conducting research relating to (but not limited to) human medical genetics, neurological disorders, stem cells as disease and development models, and regeneration biology; (iii) an excellent track record of research and publication in international peer-reviewed scientific journals in one or more related areas; (iv) an excellent track record of obtaining international and local competitive research grants; (v) extensive research and teaching experience in local or overseas higher education; and (vi) a strong track record in academic and research leadership.

Apart from teaching undergraduate, postgraduate, and general education courses, the appointee will be expected to provide leadership and develop the aforementioned related fields of research to the highest standard of excellence. Appointment will normally be made on contract basis for up to three years initially commencing June 2013 or as soon as possible thereafter, which, subject to mutual agreement, may lead to longer-term appointment or substantiation later.

Salary and Fringe Benefits

Salary will be highly competitive, commensurate with qualifications and experience. The University offers a comprehensive fringe benefit package, including medical care, a contract-end gratuity for an appointment of two years or longer, and housing benefits for eligible appointee. Further information about the University and the general terms of service for appointments is available at *http://www.per.cuhk.edu.hk*. The terms mentioned herein are for reference only and are subject to revision by the University.

Application Procedure

Please send full resume, copies of academic credentials, a publication list and/or abstracts of selected published papers, together with names, addresses and fax numbers/e-mail addresses of three referees to whom the applicants' consent has been given for their providing references (unless otherwise specified), to the Personnel Office, The Chinese University of Hong Kong, Shatin, N.T., Hong Kong (Fax: (852) 3943 1462) by the closing date. The Personal Information Collection Statement will be provided upon request. Please quote the reference number and mark 'Application - Confidential' on cover.

JP232745R

---

ience

Scientists

estigating
research.

graduate
ise. There
omics and
ical trials.
xtramural
as a very
al faculty.

or clinical

includes

programs.

NW231762R

Faculty Positions

Fudan University Shanghai Cancer Center (FUSCC), Shanghai, P.R. China

FUSCC, the first cancer hospital established in 1931 in China and now the only comprehensive cancer center in Shanghai, is undertaking a major expansion and seeking outstanding applicants for a number of faculty positions at all levels, including faculty positions of its Cancer Research Institute. We offer a highly collaborative and interactive research environment with a state-of-art research platform, tissue banking with various tumors, in vitro and in vivo imaging, and multi-disciplinary networking. Candidates with research interest in molecular cancer epidemiology, cancer signaling, mechanisms, microenvironment, cancer stem cells, system biology, translational research on molecular diagnosis and therapy are encouraged to apply. Successful applicants will be expected to develop strong, independent, and extramurally funded research programs, and to participate in graduate and medical student teaching. A highly competitive salary, start-up funds and new lab space will be provided. Applicants must have a Ph.D., M.D., or an equivalent degree, post-doctoral experience and clear evidence of research productivity and independence for junior faculty positions, and additional evidence of internationally-reputed research excellence for senior positions (full professor). The search committee will begin consideration of applications immediately and continue until positions are filled.

Applicants should submit their curriculum vitae, a list of publications within the recent 5 years, a brief summary (1-2 pages) of past accomplishments and future research plans via email (fudancancercenter@gmail.com). Application materials sent via email should be required the specified email subject. Also the applicant's position should be quoted. Detailed information regarding our center and application can also be found in http://www.shca.org.cn/eng/Careers1.asp

Contact Information:

Ms. Lu Jin
Fudan University Shanghai Cancer Center Human Resource Department
Website: http://www.shca.org.cn/eng/Careers1.asp
Add: 270 Dong'an Road, Shanghai, China Zip code: 200032
Tel: 86-21-64047456 Fax: 86-21-64175590 E-mail: fudancancercenter@gmail.com

JP232518RL

.gov

sition
Inited
se
search,

tron

the
gic
D's

TRON
NCES

NW232735R

CEPLAS
Cluster of Excellence on Plant Sciences

From Complex Traits towards Synthetic Modules

The Cluster of Excellence on Plant Sciences (CEPLAS) announces 9 faculty positions, 4 full professor (W3) and 5 tenure-track positions (W1/W2)

Innovative Strategies for Crop Improvement

Plants are the basis of all human life; they provide food and feed, medicine and drugs, and raw materials for clothing, housing and energy production. Due to global change, the role of crop plants as foundation of human civilisation becomes threatened. Innovative strategies for crop improvement are required that aim at enhancing crop productivity while minimizing resource requirements. The CEPLAS mission is to contribute new paradigms to solve urgent problems in plant performance and production through exploitation of natural variation and biodiversity. CEPLAS holds a unique position within the German Excellence Initiative as the only cluster focussing exclusively on plant sciences. It applies a new approach driven by evolutionary analysis and synthetic biology. Based on this combination paradigmatic solutions shall be provided for:

- Altering the life histories of crops
- Improving primary productivity by altering photosynthesis
- Understanding/Manipulating interactions of plants with their endophytic microflora
- Using knowledge on plant-microbe chemical communication to generate genetically enhanced plants and microbes

What We Offer

To further strengthen and complement the joint research programme of Heinrich Heine University Düsseldorf, the University of Cologne, the Max Planck Institute for Plant Breeding Research, and the Forschungszentrum Jülich, CEPLAS announces 9 faculty positions affiliated either to Heinrich Heine University Düsseldorf (HHU) or to the University of Cologne (UoC). Positions include an attractive start-up package and access to a suite of state-of-the-art research facilities. CEPLAS invites applications for:

W3 Evolutionary Plant Genomics (HHU):
The successful candidate (m/f) is expected to have a strong track record in using evolutionary and quantitative genomics approaches for the identification of gene functions that underlie complex traits.

W3 Plant Cell Biology and Biotechnology (HHU):
The successful candidate (m/f) is expected to have a strong track record focussing on the understanding and/or re-designing of fundamental plant cell functions.

W3 Plant Biochemistry (UoC):
The successful candidate (m/f) should have a strong record in plant biochemistry, preferably with a focus on secondary metabolism/metabolomics.

W3 Terrestrial Microbiology (UoC):
The successful candidate (m/f) will pursue an internationally recognised research programme with a strong focus on the community structure and function of the plant or soil microflora, preferably using genomic information.

W1/W2 Synthetic Microbiology (HHU):
The successful candidate (m/f) is expected to develop a competitive research programme focussing on the production of plant natural products in microorganism and/or on the field of genetic programming.

W1/W2 Microbial Ecological Genetics (UoC):
The successful candidate (m/f) will pursue an internationally recognised research programme on ecological principles underpinning plant-microbiome interactions. Priority will be given to candidates studying genetic and metabolic mechanisms.

W1/W2 Quantitative and Theoretical Biology (HHU):
The successful candidate (m/f) is expected to complement the experimental approaches of CEPLAS through focussing on the deduction of design principles of complex biological systems using large-scale data sets.

W1/W2 Molecular Systematics (UoC):
The successful candidate (m/f) will pursue an internationally recognised research programme in the field of trait diversification in flowering plants utilising phylogenetic and genomic information. Priority will be given to candidates studying developmental or reproductive processes.

W1/W2 Plant Genetics (HHU):
The candidate's (m/f) research programme focusses on the functional analysis of genetic programmes controlling organ development and/or plant architecture, preferably employing monocotyledonous model systems.

Conditions for employment are, in addition to the general administrative conditions in accordance with § 36 of the North Rhine-Westphalia University Act (Gesetz über die Hochschulen des Landes NRW), an aptitude for teaching, exceptional competence in research, and additional scientific achievements.

Successful candidates (m/f) are expected to actively participate in teaching in the respective study programmes. During the funding period of the excellence initiative the teaching load is reduced. Applicants must include a letter of interest, CV incl. publication list (reprints of five selected publications), statement of research and teaching interests, and contact information for at least three references. Applications should be addressed to the Dean of the Faculty of Mathematics and Natural Sciences and sent electronically to bewerbung@ceplas.de. Closing date for applications is **8 February 2013.**

The Heinrich Heine University Düsseldorf and the University of Cologne are equal opportunity employers in compliance with the German disability laws. Applications from women are expressly welcome. Women with comparable qualifications will receive particular consideration, unless another applicant displays compelling reasons to prefer this person. Applications from suitably qualified severely disabled persons or disabled persons regarded as being of equal status according to Book IX of the German Social Legal Code (SGB – Soziales Gesetzbuch) are encouraged. Disabled persons with equal qualifications will be given priority. HHU and UoC offer a Dual Career Service and are members of the Rhineland Dual Career Network (Dual Career Netzwerk Rheinland). Further information can be found at www.dualcareer-rheinland.de

HEINRICH HEINE
UNIVERSITÄT DÜSSELDORF

Universität zu Köln

Max Planck Institute
for Plant Breeding Research

JÜLICH
FORSCHUNGSZENTRUM

W232755R

UHH

DESY

Universität Hamburg

DER FORSCHUNG | DER LEHRE | DER BILDUNG

DESY and the University Hamburg invite applications for a

LEADING SCIENTIST/W3 PROFESSOR
in X-ray Nanoscience and X-ray Optics

commencing as soon as possible/**ref.no. 2163/W3**.

Applicants are expected to demonstrate research experience on an international recognized level as well as a successful track record in acquiring external project funding. The University places particular value on the quality of teaching. The University and DESY are equal opportunity employers. Both aim to increase the number of women in research and teaching and explicitly encourage women to apply. Equally qualified female applicants will receive preference in accordance with Hamburg's Higher Education Act.

Tasks and responsibilities:
The exploitation of Hamburg's highly brilliant photon facilities (FLASH, PETRA III and European XFEL) requires focused research and development in advanced X-ray Optics and imaging for nanoscience applications. We expect the candidate to contribute to the strategic development of the DESY photon science research program and to support the user operation of DESY's photon science facilities. The post holder will lecture for two hours per week at the University of Hamburg.

DESY is a world-leading research center that conceives, constructs, and operates accelerator-based research tools for exploring the structure and function of matter. Research at DESY ranges from particle and astroparticle physics to the investigation of atoms, molecules and condensed matter using novel photon facilities, including X-ray free electron lasers.

The Department of Physics at the University of Hamburg performs world-leading research in astrophysics, accelerator and particle physics, solid-state physics and nanoscience, laser physics and photon science. It hosts several cooperative research projects in nanoscience and photon science (SFB 668, SFB 925, GrK 1355).

DESY and the University of Hamburg cooperate in almost all fields within the natural sciences via their strategic partnership PIER. They also work closely together in the Center for Free Electron Laser Science (CFEL) and the Center for Structural Systems Biology (CSSB). The new cluster of excellence, the Hamburg Centre for Ultrafast Imaging (CUI), underpins Hamburg's leading position in photon science.

Requirements:
Academic qualifications and further requirements as specified in § 15 of Hamburg's Higher Education Act.

Severely disabled applicants will receive preference over equally qualified non-disabled applicants.

The successful candidate is expected to take up the position as soon as possible. For further information, please contact Prof. Dr. E. Weckert (edgar.weckert@desy.de).

Applications should include a CV, a list of publications, a summary of current and future research interests and objectives as well as a statement on teaching experience and future objectives. The deadline for applications is **21st February, 2013**. Please send applications, accompanied by the reference number **2163/W3**, to: Den Präsidenten der Universität Hamburg, Organisationsreferat, Ausschreibungsstelle, Mittelweg 177, 20148 Hamburg, Germany or via email to: UniHHAusschreibungsstelle@verw.uni-hamburg.de.

Zertifikat seit 2010 audit familiengerechte hochschule

W232725R

The University of Edinburgh

The University of Edinburgh is an exciting, vibrant, research led academic community offering opportunities to work with leading international academics whose visions are shaping tomorrow's world.

College of Science and Engineering

Schools of Engineering, Biological Sciences and Informatics

Postdoctoral Research Fellow

£30,424 - £36,298

Fixed Term: 24 months

The Flowers Consortium is an EPSRC-funded project which includes Edinburgh, Cambridge, Newcastle, King's College and Imperial College and aims to develop platform technologies for synthetic biology. We seek to fill up to four positions based in Edinburgh, each for a duration of 2 years with a possible 1 year extension. Research will be led at the Centre for Synthetic and Systems Biology (http://www.synthsys.ed.ac.uk/).

We are seeking candidates strongly motivated by research in this emergent field and keen on joining our thriving group at Edinburgh and on collaborating with our partners from the Flowers consortium. We are specifically looking for applicants for the following posts: 1) efficient and reliable DNA assembly methods, 2) quantitative characterization of biological parts, and 3) modelling techniques and design of a model-centric quantitative registry of parts.

You will have already obtained or submitted a PhD. You must clearly state in your application which post(s) you are applying for.

Apply online, view further details or browse more jobs at our website. Ref: 008287N. Closing date: 7th February 2013.

Committed to Equality and Diversity

The University of Edinburgh is a charitable body, registered in Scotland, with registration number SC005336.

www.ed.ac.uk/jobs

U232770R

UNIVERSITY OF LIVERPOOL

Faculty of Science and Engineering
School of Environmental Sciences

Postdoctoral Research Associate

£31,020 - £35,939 pa

We are seeking to recruit an innovative researcher to work on the mechanical, chemical and fluid flow effects of dehydration. You will run high-pressure, high temperature rock deformation experiments to explore the evolution of porosity, permeability and reaction rate using gypsum and serpentinite. You will characterise the resulting microstructures to ascertain what controlled that evolution.
The experimental constraints will be used to create a conceptual and numerical model for dehydration reactions in the earth and their consequences, such as earthquake generation. You will have a PhD in Geology or Geophysics and experience of experimental rock deformation. Familiarity with microstructural investigations, thermodynamics and/or numerical modeling is desirable. The post is available for 36 months.

Job Ref: R-581712/N **Closing Date: 8 February 2013**

For full details, or to request an application pack, visit www.liv.ac.uk/working/job_vacancies/ or e-mail jobs@liv.ac.uk
Please quote job ref in all enquiries.

COMMITTED TO DIVERSITY AND EQUALITY OF OPPORTUNITY

Stonewall DIVERSITY CHAMPION

POSITIVE ABOUT DISABLED PEOPLE

STANFORD
MATERIALS SCIENCE AND ENGINEERING

The Department of Materials Science and Engineering (MSE) at Stanford University invites applications for a tenure-track position at the Assistant Professor level. Under special circumstances involving exceptional academic merit, candidates at the untenured Associate Professor level may be considered. We seek applicants with significant accomplishments in materials research in its broadest sense that may include materials characterization involving structure characterization, characterization through property measurement (e.g. nano-mechanics, nano-electronics), theoretical modeling, etc. Stanford University has excellent facilities in these areas as represented by the Stanford Nanocharacterization Laboratory (SNL), the Molecular Imaging Program at Stanford (MIPS), the Stanford Nano Center (SNC), the Stanford Nanofabrication Facility (SNF), the Center for Biomedical Imaging at Stanford (CBIS) and the X-ray facilities at the Stanford Linear Accelerator Center (SLAC). The successful candidate is expected to play a major role at one of these facilities.

The successful candidate is also expected to make major contributions to large scale multidisciplinary research and to contribute significantly to leadership in Stanford's nanoscience and technology efforts which span several Departments and Schools, including faculty in the Schools of Engineering, Medicine, and Humanities and Sciences. We are looking for a motivated individual who is committed to excellence in teaching and to the mentoring of students in multidisciplinary materials research applied to energy and the environment, biomedical research and imaging, nano-mechanical or nano-electronic properties and devices, etc.

Applicants should include a summary of their educational and professional backgrounds, a current list of published work, and the names of at least three referees who may be consulted by the search committee. An indication of how the candidate's experience matches the position described above should also be given. Applicants are encouraged to write brief descriptions of their plans for future research and how those plans might be realized in a Stanford setting, as well as to submit similar statements on teaching, focusing especially on their vision of teaching to students in the Department of Materials Science and Engineering. The appointment is expected to be made during the forthcoming academic year. Please apply online at: http://mse.stanford.edu/faculty/faculty_search.html. Applications should be submitted by March 31, 2013. Questions should be directed to, Search Committee Chair, c/o Carol Scott, via electronic mail to msesearch@stanford.edu.

Professor Robert Sinclair
Chair, Department of Materials Science
and Engineering
Stanford University
Stanford, CA 94305-4034
Phone: (650) 723-1102
Fax: (650) 725-4034
Email: msesearch@stanford.edu

Stanford University is an equal opportunity employer and is committed to increasing the diversity of its faculty. It welcomes nominations of and applications from women and members of minority groups, as well as from others who would bring additional dimensions to the university's research and teaching missions.

NW232188R

MANCHESTER
1824
The University of Manchester

AstraZeneca

gsk GlaxoSmithKline

Manchester Collaborative Centre for Inflammation Research

Opportunities for Research Associates

Salaries will be within the range £29,541 - £36,298 per annum

The Manchester Collaborative Centre for Inflammation Research (MCCIR) is a new collaboration between the University of Manchester, GlaxoSmithKline and AstraZeneca to establish a world-leading translational centre for research in inflammation. The centre is directed by Professor Tracy Hussell who will be joined by Professor Daniel Davis (MCCIR Director of Research) from Imperial College London and Professor Andrew Macdonald from the University of Edinburgh. Dr Mark Travis has also joined them as a principle investigator.

Inflammation Biology — Ref: LSX-02093

Two Research Associate positions are available for 3 years to study the regulation of innate immunity in the healthy and diseased lung. Projects include: a) epigenetic regulation of innate immune activation in airway macrophages and epithelial cells in healthy, asthmatic, COPD, and post respiratory viral infection patients, b) regulation of the interaction between T cells and macrophages by the semaphorin 6D/plexin A1 interaction c) Regulation of airway macrophages by specific micro-inhibitory RNA species, modelling airway epithelial cell macrophage interactions in vitro.

Informal enquiries to Professor Hussell: tracy.hussell@manchester.ac.uk

Imaging Immune Cell Biology — Ref: LSX-02094

Up to four Research Associate positions are available for 3 years in Dan Davis's laboratory within the MCCIR (relocating from Imperial College London). Projects include (i) using super-resolution fluorescence microscopy of human immune cells to understand recognition of disease at the nano-metre scale (ii) studying Natural Killer and T cell immune synapses (iii) probing the mechanisms and functions of membrane nanotubes and (iv) studying new functions for the transfer of proteins and microRNAs between human immune cells. Details also available on the lab web-page www.dandavislab.co.uk

Informal enquiries to Professor Davis: daniel.davis@manchester.ac.uk

Intestinal Immune Responses in Health and Disease — Ref: LSX-02095

Applications are invited for a Research Associate, available for 2-3 years in the laboratory of Mark Travis. The project available will study how adaptive immunity is controlled in the intestine, focussing on intestinal dendritic cells, using both animal models and human clinical samples.

Informal enquiries to Dr. Travis: mark.travis-2@manchester.ac.uk

Please see further particulars for more information on each project.

Please apply online via the following website:

www.manchester.ac.uk/jobs

Closing date: 20 February 2013.

The University of Manchester values a diverse workforce and welcomes applications from all sections of the community.

POSITIVE ABOUT DISABLED PEOPLE

U232771R

NIGMS

Director, Division of Biomedical Technology, Bioinformatics, and Computational Biology
National Institute of General Medical Sciences
National Institutes of Health
Department of Health and Human Services

The Person: The ideal candidate will have considerable research experience demonstrating a strong understanding of computational, technological, and biological issues. In addition, candidates should possess recognized research management and leadership abilities. Candidates with substantial training and experience in computation/informatics, biomedical technology, and biomedical research will be considered. A strong understanding of the role of computation, informatics, and technology in uncovering biological principles and in advancing research on human health and disease is desired. This individual will report to the Director, National Institute of General Medical Sciences (NIGMS), but will also have access to the Director, National Institutes of Health (NIH), in coordinating activities across NIH and among federal agencies.

The Challenge: A significant challenge for the biomedical research community is the integration of the vast amount of accumulating scientific data in order to develop predictive understanding of basic biological processes. The ability to meet this challenge will be critically dependent on advances in bioinformatics and computational biology and on discovery and deployment of new technologies. The Division of Biomedical Technology, Bioinformatics, and Computational Biology (BBCB) is responsible for stimulating and funding research in these areas of importance for NIGMS. The Division supports research on bioinformatics, databases, and data mining; modeling of complex biological systems; algorithmic development and software engineering; mathematical biology, high-performance computing, molecular imaging, structural biology, and proteomics, among other areas. In addition, the Division is responsible for managing the NIH Biomedical Information Science and Technology Initiative (BISTI), an agency-wide effort to stimulate and coordinate use of computer science and technology to address problems in biology and medicine. The Division also actively collaborates with other NIH components and federal agencies in developing policies in these areas. NIGMS is seeking a leader in this field to direct the Division and the BISTI efforts, and to coordinate the work of both with other interested federal agencies and the broader scientific community. Information about the Division and BISTI is available at: **http://www.nigms.nih.gov/About/Overview/bbcb.htm** and **http://bisti.nih.gov**.

Position Requirements: Candidates must have an M.D., Ph.D., or equivalent degree in a field relevant to the position. Please see the official vacancy announcement for qualification requirements and what to submit. The position will be filled under a Title 42 excepted service appointment.

Salary/Benefits: Salary is competitive and will be commensurate with the experience of the candidate. A recruitment or relocation bonus may be available, and relocation expenses may be paid. A full package of Federal Civil Service benefits is available, including: retirement, health and life insurance, long term care insurance, leave, and a Thrift Savings Plan (401K equivalent). The successful candidate is subject to a background investigation and financial disclosure requirements.

How to Apply: The official vacancy announcement is available at: **http://www.nigms.nih.gov/About/Job_Vacancies/**

Information about the NIGMS can be found at **http://www.nigms.nih.gov/About/**.

NIGMS will begin accepting applications on **October 26, 2012**, and plans to have the position open for at least 30 days, but will not to close the application process until a candidate has been selected.

You may contact Krystal Kelly with questions about this vacancy at **kellykry@od.nih.gov** or 301-594-3827.

HHS and NIH are Equal Opportunity Employers

NW229897R

THE UNIVERSITY OF BERGEN (UiB) Norway is an internationally recognised research university with more than 14,000 students and close to 3,500 employees at six faculties. The university is located in the heart of Bergen. Our main contribution to society is excellent basic research and education across a wide range of disciplines.

UNIVERSITY OF BERGEN

Professor/Associate Professor in Marine Systematics

The Department of Biology (http://www.uib.no/bio/en) has a vacancy for a professor/associate professor in marine biosystematics. We seek broadly trained applicants who address fundamental research questions in the biology of marine benthic invertebrates. The position is intended to unite the fields of taxonomy/phylogeny, marine fauna and marine zoogeography and is particularly aimed at deep-sea communities. The successful applicant should be an active scientist with teaching expertise in several of the relevant research fields including faunistics, taxonomy, biogeography, geo-biology, phylogeny, population genetics and phylogenomics. It is also important that he/she has experience in mapping and working with dominant marine organisms through the use of classical instruments and modern technology.

Emphasis will be given to experience in project management, coordinating research activities and strategic research planning. Furthermore, the applicants should be able to document the ability to work in teams and to attract external funding.

For highly qualified applicants a start-up grant, including a post-doctoral fellowship and funds for instrumentation and running costs, may be made available for the position.

For full details and to apply please visit www.jobbnorge.no (ID no. 89283)

For further information about the position please contact
Professor Henrik Glenner; phone (+47) 55 58 45 26; e-mail henrik.glenner@bio.uib.no; or
Professor Anders Goksøyr, Head of Department; phone (+47) 55 58 44 03; e-mail anders.goksoyr@bio.uib.no.
Closing date: 10 February 2013

W232710R

Need to find the ideal candidate *fast?*

Visit

www. naturejobs .com

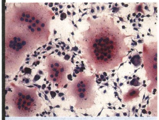
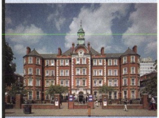
to discover how applicants can respond directly to you by email.

naturejobs
making science work

Imperial College London

Imperial College Healthcare NHS
NHS Trust

Bone and Joint Disease

Imperial College London is making major investments in the field of bone and joint disease. A new cross-Faculty venture to study bone and joint disease has been established within the Departments of Medicine and Bioengineering and consolidates existing expertise to develop new and multi-disciplinary research approaches to investigate the genetic and molecular basis of bone and joint disorders. As such we plan to make clinical and/or non-clinical academic appointments in the broad area of bone and cartilage disease.

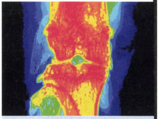
Principal investigators within the Division have substantial research funding from the Wellcome Trust, MRC, Arthritis Research UK, NIHR Comprehensive Biomedical Research Centre and funds from College-led strategic awards.

Within Imperial College London there is a broad range of clinical and non-clinical research expertise which provides an environment that is conducive to facilitating internationally competitive cutting edge research. Strong leadership, inter-group collaboration and mentorship enable senior researchers to flourish and early-stage researchers to achieve scientific independence in a supportive environment.

This initiative provides an outstanding new environment to facilitate research flexibility and enable the successful candidate to develop an independent and cross-disciplinary programme of research in skeletal science. Candidates should have broad research experience in any aspect of bone, cartilage and matrix biology.

The level of appointment (Professor, Reader, Senior Lecturer, Lecturer and Research Fellows) and the generous supporting package will be determined according to experience. Accordingly, appointments and job descriptions will be individually negotiated.

Interested individuals are encouraged to make informal enquiries to Joanna Thompson (Tel: 0208 383 3935, e-mail: joanna.thompson@imperial.ac.uk).



Valuing diversity and committed to equality of opportunity

POSITIVE ABOUT DISABLED PEOPLE

Stonewall
DIVERSITY CHAMPION

U232761R

“When trying to recruit good scientists for such a remote institutions like ours, in the Canary Is., the high visibility and wide-range exposure that Naturejobs provides has been crucial. I'm very happy to have international candidates who applied to our position from places ranging between St. Petersburg and Cincinnati. Thanks a lot!”

Eduardo Salido, MD, PhD,
Hospital Universitario de Canarias

naturejobs

Universität
Konstanz

The University of Konstanz, with its Institutional Strategy to Promote Top-Level Research, has been receiving continuous funding since 2007 within the framework of the Excellence Initiative by the German Federal and State Governments.

The Department of Biology of the Faculty of Sciences seeks to fill a

W3-Full Professorship of Cell Biology

as soon as possible.

Applicants should work and be internationally recognized in the field of Cell Biology. The applicant should integrate within one or several Research Consortia such as the Collaborative Research Initiative on “Chemical and Biological Principles of Cellular Proteostasis”, the Graduate School “Chemical Biology” or the Research Training Group RTG 1331 “Cell-based characterization of disease mechanism in tissue destruction and repair”. The Electron Microscopy Facility of the Department is linked to this Professorship of Cell Biology. The holder of the Professorship is expected to teach Cell Biology for Bachelor and Master Students within the Department of Biology.

Promising candidates will be invited to a scientific colloquium to be held in Konstanz in spring 2013.

Applications (reference number 2013/007) comprising a curriculum vitae, publication list, a list of grants and awards, details of teaching experience, as well as statements of current research topics, future research directions and interests should be sent before **February 28th, 2013** formatted into one pdf-file to Prof-2013-007@uni-konstanz.de . For further information please contact Hanns Fahlbusch, phone +49(0)7531/88-2413.

W232749R

EBERHARD KARLS
UNIVERSITÄT
TÜBINGEN

WERNER REICHARDT CENTRE
FOR INTEGRATIVE NEUROSCIENCE

The **Werner Reichardt Centre for Integrative Neuroscience (CIN)** is an interdisciplinary institution at the **Eberhard Karls University Tübingen** funded by the German Excellence Initiative program. The CIN strives to deepen our understanding of how the brain generates function and how brain diseases impair functions. It tries to make use of newly acquired insights to help people with brain disorders and to launch new mind- and brain-inspired applications in many areas of engineering and computer science. Its scientific program is guided by the conviction that progress in the understanding of brain function can only be achieved by an integrative approach spanning multiple levels of organization and pooling the knowledge of researchers from many different fields.

The Werner Reichardt Centre for Integrative Neuroscience (CIN) in Tübingen offers

Junior Research Group in Neurotechnology

In order to strengthen research on technical applications the CIN offers a junior group leader (JRG) position with tenure track option for up-and-coming young scientists with a promising track record, working in different fields of **Neurotechnology.** This might include fields such as:

- Brain machine interfaces, brain computer interfaces and neurorobotics
- Technology for sensory or motor neuroprostheses
- Rehabilitation technology
- Brain-inspired user interfaces and computing technology

We welcome additional proposals in other fields of neurotechnology. The CIN strives to increase the number of female scientists. Therefore qualified female candidates are explicitly encouraged to apply.

Junior Research Group for Neuroscientists

The CIN is offering a junior group leader (JRG) position with tenure track option for up-and-coming young scientists with a promising track record in any area of the **Neurosciences.** Since the CIN strives to increase the number of women scientists, topics well represented by qualified female candidates will be preferred. Hence we explicitly encourage female candidates to apply.

Framework:
The intended duration of the positions is for 5 years with evaluations by external experts at regular intervals. In the event of positive evaluations after 3 years, the JRG will obtain a tenure track option, which may ultimately lead to a professorship at the University of Tübingen. Start-up funds as well as substantial funding for personnel and running costs will be available, but will depend on the applicant´s qualifications and prior experience. Appointees will be full members of and active participants in the CIN, which will also provide laboratory and/or office space. The JRG leaders will be provided with opportunities to contribute to research oriented training within the framework of the CIN Graduate Training Centre and the faculties involved in the CIN will provide opportunities for the German habilitation according to established rules, if desired. According to German law, severely disabled persons with equal occupational aptitude will be given preferential consideration.

Application:
Applicants should submit a curriculum vitae, pdf files of up to 5 key publications, a statement of research achievements and future directions (not to exceed 3 pages) as well as the names and addresses of at least three referees. All documents should be submitted electronically to the Chairman of the Werner Reichardt Centre for Integrative Neuroscience Tübingen, Prof. Dr. Peter Thier, at cin@uni-tuebingen.de. For further information on the CIN see: http://www.cin.uni-tuebingen.de/. **Submission deadline is February 15th, 2013.**

CIN
Werner Reichardt
Centre for Integrative
Neuroscience

W232656R

ETH
Eidgenössische Technische Hochschule Zürich
Swiss Federal Institute of Technology Zurich

Assistant Professor (Tenure Track) of Environmental Microbiology

The Department of Environmental Systems Science (www.usys.ethz.ch) at ETH Zurich invites applications for an Assistant Professorship (Tenure Track) in Environmental Microbiology.

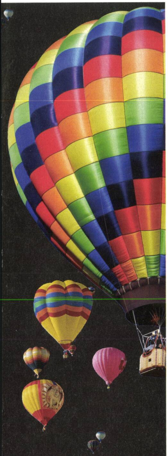
Candidates with outstanding scientific track records in any related field will be considered, but preference is given to applicants working on microbial processes in environmental systems. The new Professor should establish a world-class research group and integrate into research activities in related fields at ETH Zurich. In addition, he or she should contribute to the teaching of undergraduate level courses (German or English) and graduate level courses (English) for students of Environmental Science and Biology. The Professorship will be equipped with a generous personnel and operational budget, but it will be expected to attract further funds for research through competitive grants.

The appointment is expected to be made at the Assistant Professor level. Assistant Professorships have been established to promote the careers of younger scientists. The initial appointment is for four years with the possibility of renewal for an additional two-year period and promotion to a permanent position.

Your application should include your curriculum vitae, a list of refereed publications as well as a statement of your research and teaching interests. The letter of application should be addressed **to the President of ETH Zurich, Prof. Dr. Ralph Eichler. The closing date for applications is 30 April 2013.** ETH Zurich is an equal opportunity and affirmative action employer. In order to increase the number of women in leading academic positions, we specifically encourage women to apply. ETH Zurich is further responsive to the needs of dual career couples and qualifies as a family friendly employer. **Please apply online at www.facultyaffairs.ethz.ch.**

W232530R

Infectious career articles each week
Career advice you can't put down.
naturejobs



HIGHER & HIRE!

Post all your vacancies for FREE on the largest dedicated job board for the scientific community:

naturejobs.com

nature publishing group npg

CECAD COLOGNE
Excellent in Aging Research – Exzellent in Alternsforschung

The Faculty of Mathematics and Natural Sciences (MNF) and the Faculty of Medicine (MF) of the University of Cologne announce

THREE NEW PROFESSORSHIPS (W3)

at the Cologne Excellence Cluster on "Cellular Stress Responses in Aging-Associated Diseases" (CECAD). An attractive start-up package and access to a suite of excellent research facilities will be provided. The successful candidates are expected to further reinforce and complement the joint research efforts of the University of Cologne, the Max Planck Institutes for Biology of Ageing and for Neurological Research as well as the German Center for Neurodegenerative Diseases (DZNE).

Applicants should have a proven track record in one of the following areas

Posttranslational control of proteostasis and aging. The recruited scientist will have a strong background in biochemistry and proteome analysis and is expected to complement the expertise of existing research groups as well as to provide state-of-the-art techniques that allow qualitative and quantitative proteome analysis combined with analyses of protein modifications in different model organisms, and to act as the academic head of the CECAD Proteomics Facility. The professorship will be part of the MNF.

Lipid-dependent regulation of membrane proteins. We are seeking an expert in the field of lipid-dependent protein regulation with experience in mass spectrometry-based analysis of lipids. She/He will act as the academic head of the CECAD Lipidomics Facility to ensure continuous development of analytical procedures. This professorship will also be in the MNF.

Translational research in aging-associated diseases. We aim for efficient translation of results from promising candidate pathways into the treatment of aging-associated diseases in the CECAD research areas. Therefore, we are seeking an expert with a high-profile in translational research including the planning and execution of Phase I trials. Moreover, the candidate is expected to contribute to the expansion of the translational educational and support program of CECAD. The professor is member of the MF.

All positions are to be filled from April 2013 onward. The appointments are initially on a fixed-term basis of five years and will be turned into a permanent contract after successful review.

Applicants must include a letter of interest, a curriculum vitae, including publication list, a statement of research interest, and contact information of at least three references. Applications may be sent to the University of Cologne, c/o Pro-Rector for Research and Junior Scholars, Prof. G. Fink, Albertus-Magnus-Platz, 50674 Cologne/Germany or to gereon.fink@uni-koeln.de. The closing date for applications will be 15.02.2013.

The University of Cologne is committed to increase the percentage of women in research and teaching. Applications of women are specifically invited. In the case of equal qualifications, competence, and specific achievements, women will be considered on preferential terms within the framework of the legal possibilities.

The University of Cologne is an equal opportunity employer in compliance with the German disabilities laws. Persons with disabilities are therefore strongly encouraged to apply.

For detailed information, please visit www.cecad.uni-koeln.de

Universität zu Köln

W232765R

"*Naturejobs* has been a great way to find the right people to build our research team. In less than 1 month, I received over 40 inquiries for my advertised postdoctoral position, with many promising candidates. As a result, I have hired two very well-qualified fellows and am excited about moving our research forward with our new team."

**Dr Kathleen Martin,
Yale School of Medicine, USA**

naturejobs

naturejobs.com

Grow with us

Over the past twelve years, we've worked hard to craft naturejobs.com into an eminent career resource for scientists. In that time, with your help, we've become the largest dedicated jobs board for the scientific community worldwide.

That's why we're excited to share the new naturejobs.com – with improved functionality that lets you search, save and apply for jobs more quickly and easily.

- ✓ **Easily find jobs with our improved search**
- ✓ **Quickly save job alerts**
- ✓ **Apply faster with our CV upload feature**

Ready? Search over 10,000 vacancies to find exactly the right job for you at:
www.naturejobs.com

Follow us on:

nature publishing group npg

UNIVERSITY OF
OXFORD

www.ox.ac.uk/jobs

Project Leader & Postdoctoral Scientists Drug Discovery Programme

Weatherall Institute of Molecular Medicine

University of Oxford & the Research Complex at Harwell

An exciting research programme to discover and develop small molecule inhibitors of mutant RAS for treatment of cancer, is being initiated at the Weatherall Institute of Molecular Medicine (WIMM). This new chemical biology drug discovery programme has co-investigators based at the Research Complex at Harwell (RCaH) and the Department of Chemistry, Oxford, and is funded by a Wellcome Trust Seeding Drug Discovery Initiative.

The aim of the project is to discover small-molecule drugs that inhibit protein – protein interactions involving the RAS family of proteins for treatment of cancer. The project will adopt a highly innovative strategy using single domain antibody fragments as tools for identification of drug-like molecules.

Project Leader Ref. 106087

Lead, manage and coordinate the programme activities, with project management experience and scientific expertise in one or more of the following areas: protein production, structural biology, biophysics, and chemical biology. The medicinal chemistry component of the project will be outsourced to a Contract Research Organisation and thus you must have good communication skills in addition to experience in drug discovery, preferably within the pharmaceutical industry.

Closing date for this position: noon Friday 15 February 2013.

Postdoctoral Scientist (two posts) Ref. 106082

Postdoctoral Scientist in Structural Biology Ref. 106083

There are also positions for three Postdoctoral Scientists (two molecular biologists based at the WIMM; a protein crystallographer based at the RCaH), to carry out basic research with the small molecules developed in the medicinal chemistry component of the programme. The molecular biologists will be involved with biochemical assays and cell based assays to discover the effects inhibiting mutant RAS signalling and the structural biologist will determine the structures of small molecules in contact with RAS family member proteins and in designing improvements to compound protein interactions.

Closing date for these positions: noon Friday 8 March 2013.

Informal enquiries may be made by email to Professor Terry Rabbitts, terence.rabbitts@imm.ox.ac.uk

For more information and to apply please go **http://www.ox.ac.uk/about_the_university/jobs/research/**

You will be required to upload a CV, supporting statement as part of your online application.

Committed to equality and valuing diversity

U232772R

---

naturejobs

Need help with your career?
Ask the Expert on **naturejobs.com**

- **Browse past Q&A**
- **Get professional advice**
- **Email your questions**

Ask the Expert is located in the **Career Toolkit** section of **naturejobs.com**

THE UNIVERSITY OF HONG KONG
香 港 大 學

Founded in 1911, The University of Hong Kong is committed to the highest international standards of excellence in teaching and research, and has been at the international forefront of academic scholarship for many years. The University has a comprehensive range of study programmes and research disciplines spread across 10 faculties and about 100 sub-divisions of studies and learning. There are over 23,400 undergraduate and postgraduate students coming from 50 countries, and more than 1,200 members of academic and academic-related staff.

Teaching Consultant in Common Core Curriculum (Ref.: 201300007)

Applications are invited for appointment as Teaching Consultant in the Common Core Curriculum in the Faculty of Science, from as soon as possible, on a two-year fixed-term basis, with the possibility of renewal.

The University has established a Common Core Curriculum for all undergraduates since September 2012. The Faculty of Science offers courses in the area of Scientific and Technology Literacy to students from all ten Faculties. A senior Professor in the Faculty of Science will lead the teaching of these courses and work with the Teaching Consultants in the relevant area who will serve primarily as a tutor for small groups of students from multiple disciplines. The medium of instruction is English. Information about the Common Core Curriculum can be viewed at http://tl.hku.hk/common-core-curriculum.

Applicants should have a Master's degree, preferably in Physical Sciences and have substantial teaching experience in subjects related to general education at the university level. Those with a Ph.D. degree would have an advantage. The appointee's duties include conducting small-group tutorials for two courses per semester, marking assignments/examination scripts, preparing teaching materials, supervising student work, developing new courses, and undertaking other tasks related to science education. He/She will meet regularly with the Centre for the Enhancement of Teaching and Learning to reflect on and refine the teaching activities in Common Core courses.

A globally competitive remuneration package commensurate with the appointee's qualifications and experience will be offered. At current rates, salaries tax does not exceed 15% of gross income. The appointment will attract a contract-end gratuity and University contribution to a retirement benefits scheme, totalling up to 15% of basic salary, as well as leave, and medical benefits. Housing benefits will be provided as applicable. Please note that the University is not able to offer a relocation assistance package (including temporary University housing accommodation and a passage and baggage allowance) to a successful candidate recruited from overseas.

For enquiries about the specific job requirements, please write to Professor Sun Kwok, Dean of Science (e-mail: deansci@hku.hk). Applicants should send a completed application form, together with an up-to-date C.V. and a statement on teaching philosophy, which includes a portfolio of syllabi and descriptions of courses they have taught or co-taught by e-mail to sciappt@hku.hk. They should also arrange for submission of three references from senior academics who are familiar with their teaching approaches, skills and experience. Please indicate clearly "Ref.: 201300007 (Teaching Consultant in Common Core Curriculum)" in the subject of the email. Application forms (341/1111) can be obtained at http://www.hku.hk/apptunit/form-ext.doc. Further particulars can be obtained at http://jobs.hku.hk/. **Closes March 31, 2013**.

The University thanks applicants for their interest, but advises that only shortlisted applicants will be notified of the application result.

The University is an equal opportunity employer and is committed to a No-Smoking Policy

JP232617R

---

cefic

LRi
The Long-range Research initiative

LRI Innovative Science Award 2013

Are you a rebel thinker?
Do you have a good idea?
Make it come alive!

The European Chemical Industry Council (Cefic), in conjunction with SETAC, EUROTOX, ISES and Chemical Week, is offering a **€100,000** award to support innovative interdisciplinary research in the field of **exposure assessment to quantify human health risks.**

The award is intended for a European-based scientist with less than ten years post-doctoral experience.

Send us your proposal by March 15, 2013

For instructions and more information visit www.cefic-lri.org/awards

W232766A

REPROGRAPHICS: CC Media Group, London EC1V 8AR UK and **The Charlesworth Group,** Wakefield, UK.
PRINTED BY: William Gibbons & Sons Ltd, UK; **Publishers Press,** Lebanon Junction, Ky, USA and **Obun Printing Co. Inc,** Tokyo, Japan

a**nature**conference

Herrenhausen Symposium on Metabolic Disease

What are the most important problems to overcome to make progress in the field of metabolic disease? *Nature Medicine*, in consultation with the research community, has identified a series of roadblocks to translational research on metabolic disease. A subset of leaders in the field will attend the Herrenhausen Symposium and participate in an interdisciplinary discussion that will point in the direction of solutions to the greatest challenges that the field of metabolic disease faces.

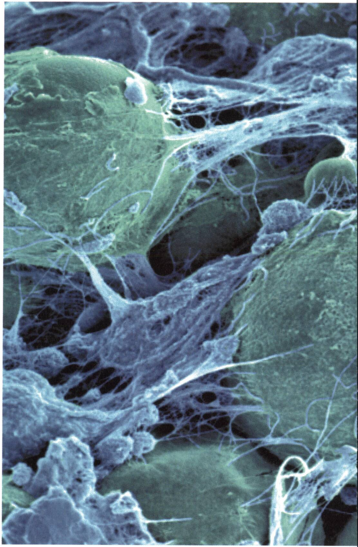
The symposium will be a closed event. In addition to the invited speakers, 10-15 attendees will be chosen to participate on the basis of their application for admission. Preference will be given to young and early-career investigators (assistant professors, lecturers, research associates, postdoctoral researchers and clinical fellows) who are currently working in related areas.

April 22-25, 2013
Hannover, Germany

ORGANIZERS:
Kevin Da Silva (*Nature Medicine*, USA)
Henrike Hartmann (Volkswagen Foundation, Germany)
Randy Levinson (*Nature Medicine*, USA)
Juan Carlos Lopez (*Nature Medicine*, USA)

CONFIRMED SPEAKERS:
Kendra Bence
Matthias Blüher
Barbara Cannon
Karine Clement
Mirela Delibegovic
Richard DiMarchi
Daniel Drucker
Jürgen Eckel
Sadaf Farooqi
Christoph Handschin
David Kelley
Anna Krook
Rohit Kulkarni
Thomas Mandrup-Poulsen
Eleftheria Maratos-Flier
Francesco Rubino
Patrick Schrauwen
Bart Staels
Scott Summers
Mathias Treier
Matthias Tschöp
Lori Zeltser
Bei Zhang

For more information and to apply, visit:
www.nature.com/natureconferences/hhs2013md



Volkswagen**Stiftung**

nature medicine

IN2327

Cold Spring Harbor Laboratory 2013 Meetings & Courses

Meetings

Meeting dates / abstracts due

From Base Pair to Body Plan – Celebrating 60 Years of DNA
February 28 - March 3 / December 14, 2012
Single Cell Analyses
March 6 - March 9 / January 4
Systems Biology: Networks
March 13 - March 16 / January 11
Computational Cell Biology: The Interplay between Models & Experimentation
March 19 - March 22 / January 11
RNA & Oligonucleotide Therapeutics
April 10 - April 13 / January 25
Synapses: From Molecules to Circuits & Behavior
April 16 - April 20 / February 1
Cancer Biology & Therapeutics
April 23 - April 27 / February 8
Telomeres & Telomerase
April 30 - May 4 / February 15
The Biology of Genomes
May 7 - May 11 / February 22
The Ubiquitin Family
May 14 - May 18 / March 1
Retroviruses
May 20 - May 25 / March 8
78th Symposium: Immunity & Tolerance
May 29 - June 3 / March 15
Wiring the Brain
July 18 - July 22 / May 3
Metabolic Signaling & Disease: From Cell to Organism
August 13 - August 17 / May 31
Eukaryotic mRNA Processing
August 20 - August 24 / June 7
Mechanisms of Eukaryotic Transcription
August 27 - August 31 / June 14
Behavior & Neurogenetics of Nonhuman Primates
September 6 - September 9 / June 21
Eukaryotic DNA Replication & Genome Maintenance
September 9 - September 13 / June 28
Microbial Pathogenesis & Host Response
September 17 - September 21 / July 8
Stem Cell Biology
September 24 - September 28 / July 12
Neurobiology of Drosophila
October 1 - October 5 / July 19
Cell Death
October 8 - October 12 / July 26
Genome Informatics
October 30 - November 2 / August 16
Cell Biology of Yeasts
November 5 - November 9 / August 23
Precision Medicine: Personal Genomes & Pharmacogenomics
November 13 - November 16 / August 30
Harnessing Immunity to Prevent & Treat Disease
November 20 - November 23 / September 6
Plant Genomes & Biotechnology: From Genes to Networks
December 4 - December 7 / September 20
Rat Genomics & Models
December 11 - December 14 / September 27

Courses

Course dates / applications due

Workshop on Leadership in Bioscience
February 22 - February 25 / January 11
Protein Purification & Characterization
April 3 - April 16 / January 31
Quantitative Imaging: From Cells to Molecules
April 3 - April 16 / January 31
Cell & Developmental Biology of Xenopus
April 5 - April 16 / January 31
Workshop on Autism Spectrum Disorders
June 5 - June 11 / March 15
Single Cell Analysis
June 5 - June 18 / March 15
Advanced Bacterial Genetics
June 5 - June 25 / March 15
Ion Channels & Synaptic Transmission
June 5 - June 25 / March 15
Mouse Development, Stem Cells & Cancer
June 5 - June 25 / March 15
Vision: A Platform for Linking Circuits, Perception & Behavior
June 12 - June 25 / March 15
Statistical Methods for Functional Genomics
June 21 - July 3 / March 15
Workshop on Pancreatic Cancer
June 26 - July 2 / March 15
Drosophila Neurobiology: Genes, Circuits & Behavior
June 28 - July 16 / March 15
Frontiers & Techniques in Plant Science
June 28 - July 18 / March 15
Advanced Techniques in Molecular Neuroscience
July 2 - July 18 / March 15
Proteomics
July 9 - July 24 / March 15
Biology & Disorders of Learning & Memory
July 20 - August 2 / March 15
Computational Cell Biology
July 23 - August 12 / March 15
Eukaryotic Gene Expression
July 23 - August 12 / March 15
Yeast Genetics & Genomics
July 23 - August 12 / March 15
Imaging Structure & Function in the Nervous System
July 24 - August 13 / March 15
Synthetic Biology
July 30 - August 12 / March 15
Cellular Biology of Addiction
August 6 - August 12 / April 15
Programming for Biology
October 14 - October 29 / July 15
X-Ray Methods in Structural Biology
October 14 - October 29 / June 15
Computational & Comparative Genomics
November 6 - November 12 / July 15
Antibody Engineering & Phage Display
November 6 - November 19 / July 15
Advanced Sequencing Technologies & Applications
November 12 - November 24 / July 15
The Genome Access Course
April 21-23, July 18-20, November 17-19

Cold Spring Harbor Laboratory
Meetings & Courses Program
1 Bungtown Road
Cold Spring Harbor, New York 11724
Phone: 516-367-8346
www.cshl.edu/meetings
meetings@cshl.edu

RISTEEN OYSTER BAY

CSHL course on an evening sail.

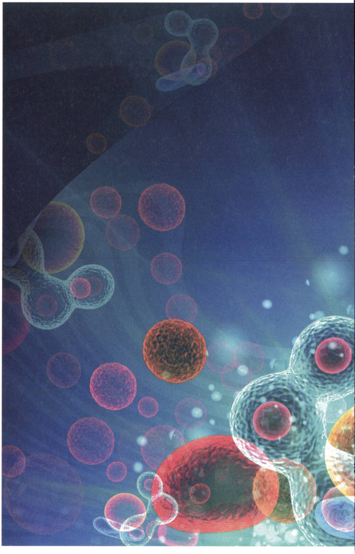
NW232733E

a nature conference

Roche - *Nature Medicine* Immunology Symposium 2013: Host Immunity to Cancer and Chronic Viral Infections

This Roche-*Nature Medicine* Symposium will bring together experts from the fields of cancer and viral immunology to highlight parallels in how the immune system fights these frequently chronic diseases. The speakers will present recent studies on host immunity and immune evasion of viruses and tumors, with the aim of improving our understanding of how to develop immune therapies that can prevent and treat these diseases.

In addition to the invited speakers, 50 participants will be selected by application. Registration is complimentary on acceptance.



April 28-30, 2013
Roche Forum, Buonas, Switzerland

SCIENTIFIC ORGANIZERS:

George Coukos (University of Lausanne, Switzerland)
Alison Farrell (*Nature Medicine*, USA)
Janet Hammond (Roche pRED, Switzerland)
Hy Levitsky (Roche pRED, Switzerland)
Juan Carlos Lopez (*Nature Medicine*, USA)

CONFIRMED SPEAKERS:

Matthew Albert
Vincenzo Bronte
Frank Carbone
Douglas Fearon
Jacques Fellay
Olivera Finn
Philip Greenberg
Wayne Hancock
Guido Kroemer
Thomas Kupper
Stanley Lemon
Sharon Lewin
Cornelius Melief
Drew Pardoll
Bali Pulendran
Carola Ries
Cornelia Liu Trimble
Christopher Walker

The application and abstract submission deadline is **February 21, 2013**. Three presented posters will be selected to receive an award.

For more information and to apply, visit:
www.nature.com/natureconferences/rnms2013

Roche

nature medicine

CSH Asia

2013

Cold Spring Harbor Asia Meetings

中國·蘇州
Suzhou, China

Suzhou Dushu Lake Conference Center
No.299 Qiyue Road SIP/ Suzhou, Jiangsu, China
Phone: +86 512 6272 9029 Fax: +86 512 6272 9028
E-mail: meetings@csh-asia.org

Mechanisms and Functions of Non-apoptotic Cell Death April 15-19
Junying Yuan, Jiahuai Han, Masayuki Miura, Peter Vandenabeele

Francis Crick Symposium on Neuroscience May 6-10
Nancy Ip, Liqun Luo, Karel Svoboda

Membrane Protein Structure and Function May 13-17
Carola Hunte, Brian Kobilka, Robert Nakamoto, Nieng Yan

Metabolism, Obesity and Obesity-associated Diseases May 20-24
David James, Karen Reue, Baoliang Song

Vaccine Design June 3-7
Michael Good, Yiming Shao, Linqi Zhang

Plant Cell and Developmental Biology June 17 -21
Tetsuya Higashiyama, Bo Liu, Hong Ma, Zhenbiao Yang

Yersinia 11 June 24-28
Ruifu Yang, Elisabeth Carniel, Paul Keim, Andrey Anisimov

New Advances in Optical Imaging of Live Cells and Organisms August 20-23
Atsushi Miyawaki, Xiaoliang Sunney Xie, Rafael Yuste, Xiaowei Zhuang

Cell Signaling in Metabolism, Inflammation and Cancer September 2-6
Michael Karin, An-Ning Lin, Zheng-Gang Liu, David Wallach

Molecular Basis of Aging and Disease September 9-13
Adam Antebi, Jing-Dong Jackie Han, Brian Kennedy, Jan Vijg

Frontiers in Bioinformatics and Computational Biology September 23-27
Madan Babu, Luonan Chen, Keith Dunker, Satoru Miyano, Dong Xu

Genetic, Genomic and Translational Studies of Human Leukemia October 7-11
Scott Armstrong, Yushiaki Ito, Paul Liu, Pier-Paolo Pandolfi

CSHA/ISSCR Joint Meeting on Stem Cells in Science and Medicine October 14-18
Andrew Elefanty, Hongkui Deng, Ronald McKay, Huck-Hui Ng, Richard Young

Development, Function and Disease of Neural Circuits October 21-25
Barry Dickson, Zhigang He, Hitoshi Okamoto, Yimin Zou

Tumor Immunology and Immunotherapy October 28-November 1
Xuetao Cao, Olivera Finn, Cornelius Melief

Nuclear Receptors and Diseases November 4-8
Hueng-Sik Choi, David Moore, Kristina Schoonjans, Nanping Wang

Bacterial Infection and Host Defense November 18-22
Sam Miller, Craig Roy, Feng Shao, Gisou van der Goot

For the most updated information,
please visit our website at
www.csh-asia.org

Built in 825 A.D., 19.8km away from CSH Asia, 30 minutes by driving, Shantang Street is located in the northwest corner of the old town of Suzhou. It was one of the few very early spots where commerce was spontaneously budded after more than two thousands years of peasantry-based economy within China. This area also symbolizes the climax of the economical prosperous of the Chinese society in the Ming and Qing dynasties. Strolling on the pebble-paved Shantang street inevitably provides a glimpse of the past glory of old Suzhou / Photographer: Bifeng LI

EROS FOR ANNABELLE

Love's labour.

BY IAN WHATES

I never liked Annabelle. During one particularly heated exchange, Josh snapped that my animosity was born of jealousy. That hurt; in part because he probably had a point. Annabelle was gorgeous, intelligent, witty, charming, vindictive, manipulative, ambitious and calculating. I could see all of that. Josh, naturally, could see only half of it. What mother wouldn't be envious of such a woman … particularly when her son's heart had been stolen away? And he was all I had left.

Annabelle latched onto Josh because he was a genius, and abandoned him when it became apparent that his brilliance would never bring the commercial, the *financial* success she craved. At first she worked hard on his behalf, I'll grant her that much; but, despite all the doors that her striking looks and social connections could open, Josh's talent proved too abstract and too elusive.

JACEY

Given the uncertainty of the still-emerging economy, we all longed for some security; but Annabelle more than most. Inevitably she caught the eye of someone wealthy and influential, transferring her affections in a flash and slipping into this new relationship as if it were a natural progression: social evolution in action. Josh was devastated, as was I, by proxy. I hated Annabelle with a vengeance, but I hated what her loss did to my son even more.

Faced with an impossible choice, I chose the unthinkable: encouraging him to try to win her back.

Of course, it was never going to be that easy. Did I mention that Annabelle was vindictive? She set him a challenge. Presumably she found the prospect of his scrabbling around in a desperate, futile attempt to impress her amusing. It wasn't because she expected or wanted him to succeed, that much is certain. After all, the challenge she set him was impossible, but he wouldn't refuse — *couldn't* refuse.

Eros: that was what she wanted. Annabelle had worked in London before the Downfall and kept eulogizing the place: Shangri-La with added fairy dust to hear her speak of it. To prove his love for her, she demanded that my son recreate a tiny piece of her precious London: the corner of Piccadilly Circus where the famous statue of Eros, God of Love, once stood. She gave him until New Year's Eve. Fail, and he would lose her forever.

NATURE.COM
Follow Futures:
@NatureFutures
go.nature.com/mtoodm

The whole thing doubtless struck her as deliciously ironic: love for love; and at New Year too, which was, of course, the last great celebration; the last joyous time that anyone could remember from the Old World. Downfall had come in January — the fiery destruction of the cities and the collapse of the Internet: the end of a civilization we'd all so blithely taken for granted.

No one knew who started it and no one cared, not any more. People's horizons had narrowed dramatically. They had to. It was that or perish. The 'Downfall Virus' — a misnomer as Josh kept telling me — destroyed the Internet and rendered every computer useless. It wasn't really a single virus, Josh claimed, but a multitude of nanotech that kept replicating, adapting and mutating. Everything seemed to be infected: even isolated computers that had never been linked to the net would crash and go into meltdown without warning. I didn't understand how it worked, but Josh did. That was his genius.

"All the information that existed on the net, it's not really gone," he once told me. "It's out there somewhere, but we're being kept from it. The Downfall viruses are blocking our access, that's *their* purpose; and mine is to defeat them, to find a way of circumventing Doomsday and regaining what we've lost — as much of it as possible, at any rate."

Success would ensure that he was set for life. Annabelle knew that — had *counted* on that — until she lost faith. And now she taunted him, insisting he produce instant triumph to order. Only with a computer could he hope to build the 3D rendering of Eros she required. He tried, oh how he tried. No one could have strived harder. He worked like a man possessed, with reckless disregard for his own physical and mental health. Sleep, food, hygiene, they would all have been neglected entirely if not for the attentions of a worried mother and the limits of human endurance.

Even so, he failed. In frustration and the anguished conviction that he was close … Christmas and New Year came and went, as did Annabelle. An exhausted Josh had a breakdown. I feared for him then, feared that he would never recover. But in time he did, thanks to the patience and love of Jasmine, who is everything that Annabelle could never be. Did I mention that my son is a genius? He returned to work, succeeding as I always knew he would.

Doomsday is still with us but it's manageable now. We're learning to cope, thanks to Josh. And, drawing on a rudimentary, partially reclaimed Internet, he finally built his Eros, inspired by someone who is worthy of such a tribute, a woman who is everything that a mother could hope for. He's famous now, and happy, and wealthy beyond Annabelle's wildest dreams.

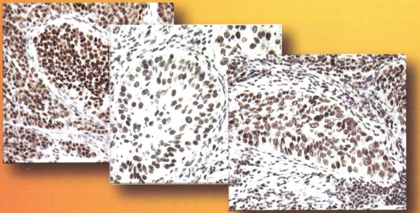
Want to hear the funny part? The statue which that bitch drove Josh to recreate was never really of Eros at all, though that's what everyone called it back in the Old World. It was actually of Anteros, Eros' brother. They both numbered among the Erotes, the ancient Greeks' Gods of Love, but Anteros was a very specific god. He was the Avenger of Unrequited Love. It would seem that Annabelle knew less about irony than she supposed. ■

Ian Whates *is the author of the* Noise *books (Solaris) and the* City of 100 Rows *trilogy (Angry Robot). His second collection,* Growing Pains, *is due from PS Publishing in April 2013.*

UltraMAB™

The Ultra Specific Antibodies

Validated against >10,000 human antigens



Trial Size UltraMABs Available!
OriGene.com/UltraMAB

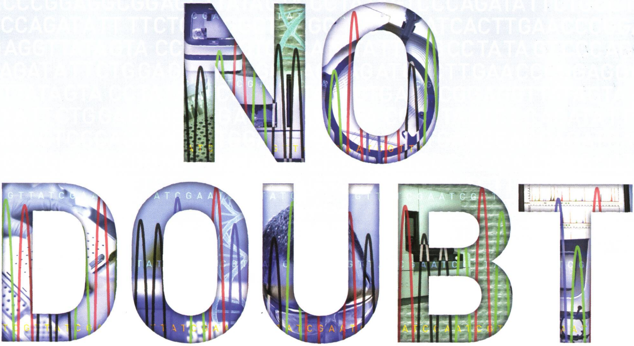
ltraMAB = Minimal Cross Reactivity!

ltraMAB is the only class of antibodies rigorously
alidated for mono-specificity. Each UltraMAB has
een tested against >10,000 human protein antigens
rrayed on a high-density protein chip. The
nprecedented specificity test ensures the superior
erformance characteristics of minimal cross-reactivity
nd exclusive specificity.

ltraMABs currently available -
ER2, Beta Catenin, P53, ERCC1, Keratin 8, CD2, etc.

ORIGENE
Your Gene Company
origene.com/UltraMAB

The products represented on this page are for research use only. These products are not intended to be used for therapeutic or diagnostic purposes in humans or animals.



applied biosystems®
by life technologies™

SANGER SEQUENCING.

We love it. Build it. Support it.
Service it. Know it. Trust it. Live it.

See the widest range of applications from the most-cited provider at **lifetechnologies.com/sanger**

life technologies™

For research use only. Not intended for any animal or human therapeutic or diagnostic use. ©2012 Life Technologies Corporation. All rights reserved. The trademarks mentioned herein are the property of Life Technologies Corporation or their respective owners. CO25454 0812